

SEARCH REQUEST FORM

Requestor's
Name: _____Serial
Number: _____

Date: _____

Phone: _____

Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLYDate completed: 06-02-01Searcher: Beverly, @ 4994Terminal time: 20

Elapsed time: _____

CPU time: _____

Total time: 25

Number of Searches: _____

Number of Databases: 1**Search Site**

____ STIC

____ CM-1

____ Pre-S

Type of Search

____ N.A. Sequence

____ A.A. Sequence

____ Structure

____ Bibliographic

Vendors

____ IG

____ STN

____ Dialog

____ APS

____ Geninfo

____ SDC

____ DARC/Questel

☒ Other CGN

QY 363 EKLOERLAKVGVAVIHHGAATEEMKEKDRVEDALNATRAAVEGIVPGGTAFAVRS 422
 DB 361 ekIqerIakIagvavikyaateveIkerkrvedalnatraaveegivpggyallra 420
 QY 423 IKVLDDIKPADDELALGNIIRSRLEPLRQIAANAGYGSIVKEVREPKGFCFNAS 482
 DB 421 apaldIkIktengdeatgvnIvIraIeaPIrIgaenaIegsvvekvkseag-gynaat 479
 QY 483 GEYEDLIKAGVIDPKKVTPTALONASVSLITTECAIAEKPEPKOMP--MPC--GGM 538
 DB 480 geyvdmIaagIIdpKvtIrsalIqnaasvasImItteavvdkpe-keaaPgmgmgm 538
 QY 539 GGMGM 544
 DB 539 gmgmgm 544
 ULT 3
 ID AAY23909 standard: Protein: 549 AA.
 AC AAY23909:
 DT 22-SEP-1999 (first entry)
 DE Amino acid sequence of a heat shock protein.
 XX Heat shock protein: Hsp; immune response; immunological carrier;
 KW cancer control; tumour; sarcoma; cancer; gene therapy.
 XX Legionella pneumophila.
 OS WO9335270-A1.
 PN 15-JUL-1999.
 PD 29-DEC-1998: 98WO-CA01203.
 PF 31-DEC-1997: 97US-0001737.
 PR (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
 XX Mlzen L, Wisniewski J;
 PI WPI: 1999-430397/36.
 XX New nucleic acid encoding heat shock protein-60 from Streptococcus,
 PT useful in vaccines, as carriers for other immunogens, as anticancer
 PT agents and for diagnosis
 XX Disclosure: Fig 10A-E; 176pp; English.
 PS AAY23905-30 represent heat shock proteins (Hsps). The specification
 CC describes Streptococcal Hsps, designated Hsp60. These proteins, their
 CC fragments, variants and fusion proteins, are used to elicit or enhance
 CC an immune response against Streptococcus, and to elicit a similar
 CC response to a target antigen fused to the protein. Unlike other
 CC immunological carriers, Hsp60 proteins are not immunosuppressive so
 CC provide an increased response to any conjugated or fused antigen. Also,
 CC where used for cancer control, they lack the side effects associated
 CC with endotoxins. They can also be used to detect specific antibodies
 CC and in treatment or prevention of tumours (e.g. sarcoma or cancers of
 CC breast, ovary, prostate, lung, pancreas or liver). The Hsp60
 CC polynucleotide is used for recombinant production of the protein, as
 CC a source of primers and probes for detecting streptococci in standard
 CC hybridization/amplification assays, and therapeutically in gene
 CC therapy vectors.
 CC
 XX Sequence 549 AA:
 SO

Query Match 68.2%: Score 1848; DB 20: Length 549;
 Best Local Similarity 67.3%: Pred. No. 5,9e-114;

Matches 367; Conservative 80; Mismatches 96; Indels 2; Gaps 1:
 QY 3 SKELFDKARREKLSRGVDKLANAVKVTLPGRNVIEKSGSPVITKGVSVAKTEL 62
 DB 2 akelIftfgdardIqmiIagvnaIadevqcmPgrrnvIeksgIgapIctvctdgySvakeIef 61
 QY 63 EDKFEHQAWKVEKAPKTSIDAGDGTATVLAQIYREGVKYLAAGNPAKIRGIDK 122
 DB 62 ehIrfmmgawmKveasKtsIdagdgdtIatvIarsIveghkavaagmpndIkrgIdk 121
 QY 123 AVAATYKELSDITTPRDKETIAOVCTISANDDTIGNIAAMKVKVCGVITVEAAG 182
 DB 122 avIavtkkIqamSkpckdSkalaqvtIsandsaIgaIlaaemKvKgyvIvIvedgng 181
 QY 183 LETTLDVVEGKMFDRGYSIPYFVTNPEKMWCELDNPYILCNEKITSMKDMPILEQVAK 242
 DB 182 IeneIsVvegmqfdrgyIsipyfIlnqgmScelchpIilIvdkvSiremsIvIevgK 241
 QY 243 VNRPLIIADEVGEALATLVNKLGAQVAVAKAGFERKKALEDIALITGGEAIF 302
 DB 242 sgIrpIIIAedvgealeclIvnmrgIvKvcavKagfgdrkkaIqdaIilIkgyIs 301
 QY 303 EDRGIRLENVSLSLCTAKRVVIDKENTTVIGAGKSEDIKARVKQIRQIERTSSDYR 362
 DB 302 eelgKalegaleclIdgsakrIvvtKentIIdgqKaleInarIqIraQmeetItdydr 361
 QY 363 EKLOERLAKVGVAVIHHGAATEEMKEKDRVEDALNATRAAVEGIVPGGTAFAVRS 422
 DB 362 ekIqerIakIagvavikyaateveIkerkrvedalnatraaveegivpggyallra 421
 QY 423 IKVLDDIKPADDELALGNIIRSRLEPLRQIAANAGYGSIVKEVREPKGFCFNAS 482
 DB 422 gkalIdIkgnndqnmngInIrralespmrqtItnagSvsvvnkvaehkdhgYfnaat 481
 QY 483 GEYEDLIKAGVIDPKKVTPTALONASVSLITTECAIAEKPEPKOMPMPGCGMG 542
 DB 482 geygdmvengIIdpKvtIrsalIqnaasvasImIttecmvadI--pkkeegvagdmgng 539
 QY 543 GMDGM 547
 DB 540 gmgmgm 544
 RESULT 4
 ID AAB69060 standard: Protein: 547 AA.
 AC AAB69060:
 DT 18-APR-2001 (first entry)
 DE Pseudomonas aeruginosa groEL protein sequence.
 XX Pseudomonas aeruginosa; chitinase; groEL; chIA; antigen; vaccine;
 KW diagnosis; detection; infection; immune response.
 XX Pseudomonas aeruginosa.
 OS WO200102577-A1.
 PN 11-JAN-2001.
 PD 03-JUL-2000: 2000HO-GH02554.
 PF 01-JUL-1999: 99GB-0015419.
 PR (PROV-) PROVALIS UK LTD.
 XX Smith CJ, Thompson SE, Smith MW, Peek K, Sizer PJH,
 PI WPI: 2001-080988/09.
 DR N-PSDB: AAF32452.
 XX

PT Antigenic Pseudomonas aeruginosa proteins, useful in the detection
 and/or diagnosis of *P. aeruginosa* infections and for producing vaccines
 against *P. aeruginosa*.

PS Claim 3: Fig 25: 129pp; English.

CC The present invention describes antigenic Pseudomonas aeruginosa
 CC proteins (P1). The *P. aeruginosa* proteins have antibacterial activity
 CC and can be used in vaccines and as antagonists. The proteins or their
 CC fragments, or antibodies are useful in the detection and/or diagnosis
 CC of *P. aeruginosa*. They are also useful for producing a vaccine and
 CC inducing an immune response against *P. aeruginosa* infection. An agent
 CC capable of antagonizing, inhibiting or otherwise interfering with the
 CC function or expression of P1 are useful in the manufacture of a
 CC medicament for the treatment or prophylaxis of *P. aeruginosa* infections.
 CC The present sequence represents a specifically claimed *P. aeruginosa*
 CC GroEL protein sequence from the present invention.

XX Sequence 547 AA:

Query Match 67.4%; Score 1825.5; DB 22: Length 547;

Best Local Similarity 66.8%; Pred. No. 1.8e-112; Mismatches 102; Indels 5; Gaps 2;

atches 367; Conservative 75; Mismatches 102; Indels 5; Gaps 2;

1 MASEKLEFADAKAREKLSKGVDLNANAVKVTLCGPKGRNVIEKSFSPVITKDGVSVAKEI 60

1 MAKEVKIGDSARKKMLVGVNLADAVKATLGPGRNVIEKSFSPVITKDGVSVAKEI 60

61 ELEKFEKMGQMKVAPKTSIDAGDCTTATVLAQAIYREGVKVLAAGRNPAIKRGI 120

61 ELKDKFEMNGEQLVVDASKANDAGDGTATVLAQAIYREGVKVLAAGRNPAIKRGI 120

121 DKAVAAVTKELSDITKPTROKKEIAOVGTISANSSTTIGNIAEMAKVKGCVITVEEA 180

121 DKAVAAVTKELSDITKPTROKKEIAOVGTISANSSTTIGNIAEMAKVKGCVITVEEA 180

181 KGLFTLDVVEGKMFDRGYLSPYFVNPEKVCEDLNPIYILCNKIKTSMKMDLPILROY 240

181 KGLFTLDVVEGKMFDRGYLSPYFVNPEKVCEDLNPIYILCNKIKTSMKMDLPILROY 240

241 AKVAPRLIITAEVDEGALATLVNKLKALQVAVKAPGFGERRKAMLEDAITLTCGEA 300

241 AKVAPRLIITAEVDEGALATLVNKLKALQVAVKAPGFGERRKAMLEDAITLTCGEA 300

421 AKAPRPIIIVAEDEVGALATLVNKLKALQVAVKAPGFGERRKAMLEDAITLTCGEA 360

421 AKAPRPIIIVAEDEVGALATLVNKLKALQVAVKAPGFGERRKAMLEDAITLTCGEA 360

301 IFEDRGKLENVSLSTGTAKRVYIDKENTTIVDAGKSEDIKARVQIRAOIEETSSDY 360

301 IFEDRGKLENVSLSTGTAKRVYIDKENTTIVDAGKSEDIKARVQIRAOIEETSSDY 360

301 ISEEVGISLEGATLHNGAKRVYIDKENTTIVDAGKSEDIKARVQIRAOIEETSSDY 360

361 DREKLOERLAKLVGVAVIHVGAATETEMKEKKDRVEDALNATRAAVEGIVPGGTAFAV 420

361 DREKLOERLAKLVGVAVIHVGAATETEMKEKKDRVEDALNATRAAVEGIVPGGTAFAV 420

421 RSIKVLDDIKPADDELACGLNIIRSLSEPLRQIAANAGYEGSIVKVRPEPKDGFNA 480

421 RSIKVLDDIKPADDELACGLNIIRSLSEPLRQIAANAGYEGSIVKVRPEPKDGFNA 480

481 ASCEYEDLITAGVYIDPKRVIRIALONASVASTLTTECAIAEKPEKKMPHGG--GM 538

481 ASCEYEDLITAGVYIDPKRVIRIALONASVASTLTTECAIAEKPEKKMPHGG--GM 538

539 GCMGCMGDM 547

538 GCMGCMGDM 546

538 GCMGCMGDM 546

538 GCMGCMGDM 546

RESULT 5

ID AAR67381 standard: Protein: 547 AA.

AC AAR67381:

XX 22-JUN-1995 (first entry)

XX L. pneumophila HtpB gene product.

XX Urease; Immunogen; vaccine; diagnostic; t

XX GroEL-like protein; Helicobacter felis.

XX Legionella pneumophila.

XX WO9426901-A.

XX 24-NOV-1994.

XX 19-MAY-1994; 94WO-EP01625.

XX 19-MAY-1993; 93EP-0401309.

XX 19-NOV-1993; 93WO-EP03259.

XX (INRM) INST NAT SANTE & RECH MEDICALE.

XX (INSP) INST PASTEUR.

XX Ferrero R, Labigne A, Suerbaum S, Thiberge J;

XX WPI: 1995-006797/01.

XX DNA from Helicobacter pylori and Helicobacter felis - used to

XX develop prods. for detection, treatment and prevention of

XX Helicobacter infection

XX Disclosure: Fig. 7A(1-11); 168pp; English.

XX The sequence of the Helicobacter pylori heat shock protein A

XX (given in AAR67381) was compared to that of other GroEL-like

XX proteins from Legionella pneumophila (AAR67381), Escherichia coli

XX (AAR67382), Chlamydia psittaci (AAR67383), Mycobacterium leprae

XX (AAR67384) and human mitochondrial protein P1 (AAR67385), and regions

XX of homology were identified.

XX Sequence 547 AA:

Query Match 67.2%; Score 1820.5; DB 16: Length 547;

Best Local Similarity 66.6%; Pred. No. 3.8e-112; Mismatches 97; Indels 3; Gaps 2;

Matches 363; Conservative 82; Mismatches 97; Indels 3; Gaps 2;

3 SKELFPAKAREKLSKGVDLNANAVKVTLCGPKGRNVIEKSFSPVITKDGVSVAKEI 62

3 SKELFPAKAREKLSKGVDLNANAVKVTLCGPKGRNVIEKSFSPVITKDGVSVAKEI 62

2 AKELIFGDDARIGMLAGVNLADAVGVTMPGRNVIEKSFSPVITKDGVSVAKEI 61

2 AKELIFGDDARIGMLAGVNLADAVGVTMPGRNVIEKSFSPVITKDGVSVAKEI 61

63 EDPEENMGAMVKEVAPKTSIDAGDCTTATVLAQAIYREGVKVLAAGRNPAIKRGI 122

63 EDPEENMGAMVKEVAPKTSIDAGDCTTATVLAQAIYREGVKVLAAGRNPAIKRGI 122

62 ENHFMNGAMVKEVAPKTSIDAGDCTTATVLAQAIYREGVKVLAAGRNPAIKRGI 121

123 AVAAVTKELSDITKPTROKKEIAOVGTISANSSTTIGNIAEMAKVKGCVITVEEA 182

123 AVAAVTKELSDITKPTROKKEIAOVGTISANSSTTIGNIAEMAKVKGCVITVEEA 182

122 AVAAVTKELSDITKPTROKKEIAOVGTISANSSTTIGNIAEMAKVKGCVITVEEA 181

183 LETTLDVVEGKMFDRGYLSPYFVNPEKVCEDLNPIYILCNKIKTSMKMDLPILROY 242

183 LETTLDVVEGKMFDRGYLSPYFVNPEKVCEDLNPIYILCNKIKTSMKMDLPILROY 242

182 LENELYVEGMDIRGYLSPYFVNPEKVCEDLNPIYILCNKIKTSMKMDLPILROY 241

WILKINSON MC;

DB 421 gkaldeIskgdnddgmngInIlIrralespmrqlvtnaagyaasvvnkhvachkdnfygfnat 480
 QY 483 GYEEDILKAGVIDPKKVTIRIALQNAASVASLLTTECAIAEKPEPKKDMPPGCGMG 542
 DB 481 geygdmvengIldpckvtrmalqnaasvasImlttecwmadl--pkkeegvgagdgmg 538
 QY 543 GMDGM 547
 DB 539 gmggm 543

RESULT 6

AAAB69061
 ID AAB69061 standard; Protein; 574 AA.

XX

AAB69061;

18-APR-2001 (first entry)

XX

pTrrHisB expression vector groEL fusion protein sequence.

DE

XX Pseudomonas aeruginosa; chitinase; groEL; chiA; antigen; vaccine;
 KW diagnosis; detection; infection; immune response.

XX

Pseudomonas aeruginosa.

OS

Synthetic.

XX

WO200102577-A1.

XX

11-JAN-2001.

XX

03-JUL-2000; 2000WO-CB02554.

XX

01-JUL-1999; 99GB-0015419.

XX

(PROV-) PROVALLIS UK LTD.

XX

Smith CJ, Thompson SE, Smith MW, Peek K, Sizer PJH, Wilkinson MC;

XX

WPI; 2001-080988/09.

XX

N-PSDB; AAF32453.

XX

Antigenic Pseudomonas aeruginosa proteins, useful in the detection
 and/or diagnosis of P. aeruginosa infections and for producing vaccines
 against P. aeruginosa -

PS

Claim 3; Fig 26; 129pp; English.

XX

XX The present invention describes antigenic Pseudomonas aeruginosa
 CC proteins (P1). The P. aeruginosa proteins have antibacterial activity
 CC and can be used in vaccines and as antagonists. The proteins or their
 CC fragments, or antibodies are useful in the detection and/or diagnosis
 CC of P. aeruginosa. They are also useful for producing a vaccine and
 CC inducing an immune response against P. aeruginosa infection. An agent
 CC capable of antagonising, inhibiting or otherwise interfering with the
 CC function or expression of P1 are useful in the manufacture of a
 CC medicament for the treatment or prophylaxis of P. aeruginosa infections.
 CC The present sequence represents a specifically claimed pTrrHisB
 CC expression vector groEL fusion protein sequence from the present
 CC invention.

XX

Sequence 574 AA;

SQ

Query Match

Best Local Similarity 67.2%; Score 1819; DB 22; Length 574;
 Matches 365; Conservative 75; Mismatches 103; Indels 4; Gaps 1;

OY 1 MASKELIFDAKAREKLSRGVDKLANAVKVTIAGPKGRNVVLEKSGSPVITKDCGSYAKEL 60
 DB 31 makekvkfgdsarkhnlvgvniadavkalipkgrnvldksfgapltlkdgvsakel 90

OY 61 ELEDKFNMGAMVKEVAPKTSIDAGDGTTFATVLQAIYREGVQLVAGRNPMATIKRGI 120
 DB 91 elkdKfemmgagqvkdvaskandaagdgcttatvtagalvneglkvaagmmpmlkrgi 150
 OY 121 DRAVAAVTKELSDITKPTBDQKEIAQVGTISANSDTJTJNIIAEMAKVKGCVITVEBA 180
 DB 151 dkatvalvaqklkelakpcadtkalagvgtlsansdesiglllaamekvgegtlvveag 210
 OY 181 KGLETTLDVVEGKFKDRGLSPFYVNPPEKMYCELDNPIYLCNEKKYTSKKMLPLBOY 240
 DB 211 sglelsvvegmqfdrylspytvnpkpdmaaeldspililvdkklsnrlremlpvleav 270
 OY 241 AKVNRPLLIIEDEVEGEALATLVNKLRCALOVAVAKAFGFRKAMLEDIATILGGEA 300
 DB 271 akagrpilivaeevegealatlvmnmrglvkvaakpdyfgortkamigdalitgylv 330
 OY 301 IFEDRGICLENVSLSSLTGAKRVYIDKENTTVIVDAGKSEDIKARVQIARAQIEHTSSDY 360
 DB 331 lseevglstegatlgnakrvvlnkentltdggvgadlearvldqrlkqteetsty 390
 OY 361 DREKLOERLAKLVGVAVIHYGATETEMKEKKDRVEDALNMTAAVEGIVPGGTAFV 420
 DB 391 dreklqerlaklvagvavikvgatevemkekkravedalhatraavegvvpggyvalv 450
 OY 421 RSIKVLDDIKPADDDLAGLNIIRSLIEPLRQIAANAGEGSIYVEKXREPKDGFUNA 480
 DB 451 raigaleglkgdneeqviallrravesplrqivanaagdepsvvdkvkgpsgnyfna 510
 OY 481 ASGEYEDLTKAGVIDPKKVTIRIALQNAASVASLLTTECAIAEKPEPKKDMPPGCGMG 540
 DB 511 atgyvgdmlemgldpdkvtrsalqaaalslgmltlemvaeivedkpm-----gmpmd 566
 OY 541 MCGMDGM 547
 DB 567 mgmggm 573

RESULT 7

AAAY23915
 ID AAY23915 standard; Protein; 545 AA.

XX

AAAY23915;

XX

22-SEP-1999 (first entry)

XX

Amino acid sequence of a heat shock protein.

DE

KW Heat shock protein; Hsp; immune response; immunological carrier;
 KW cancer control; tumour; sarcoma; cancer; gene therapy.

XX

Neisseria meningitidis.

OS

WO9935270-A1.

XX

15-JUL-1999.

XX

29-DEC-1998; 98WO-CA01203.

XX

31-DEC-1997; 97US-0001737.

XX

(STRE-) STRESSGEN BIOTECHNOLOGIES CORP.

XX

Mizzen L, Wisniewski J;

XX

WPI; 1999-430397/36.

XX

PT New nucleic acid encoding heat shock protein-60 from Streptococcus,
 PT useful in vaccines, as carriers for other immunogens, as anticancer
 PT agents and for diagnosis

XX

Disclosure; Fig 10A-E; 176pp; English.

XX

AAAY23905-30 represent heat shock proteins (hsps). The specification

QY	241	AKRRPILIIIAEDVEGALATTLVNNKLRGLQVVAAKAGFGFRRRAAMEIDAITLTGGA	300
Db	241	AKASRPILIIIAEDVEGALATTLVNNKLRGLQVVAAKAGFGFRRRAAMEIDAITLTGTV	300
QY	301	IFEDRGITKLENVSSISGTAKRRVVDKENTTVDGAGKSESDIKARVKQIRAOIEFTSSDY	360
Db	301	ISEEVLGSLKELATLDDYGQKRLKELGKENTLIDGFGDAAGLEARVAELRQGLKELATSDY	360
QY	361	DREKLERLAKLVGCAVIVHVGAAETTEMKEKKDRVEDALNATRAAVERGIVPGGCTAFV	420
Db	361	DKEKLGKRVAKLVGCAVIVHVGAAETTEMKEKKDRVEDALNATRAAVERGIVPGGCTAFV	420
QY	421	RSIVLDDIRPADDEDLAAGLNIIRSLSEPLRLQIAANAGYEGSIVYEKIREKDFGRFA	480
Db	421	RATATLNLTLGNADGDAQVGLVTRAVESPLIRIVANAGGEPSSVNVKLEGLKQNYGNA	480
QY	481	ASGEEDLIKAGVDPKKVTRIALQNAASVASLLTTECAIAEKPEPKDMPGGCMGC	540
Db	481	GSGEYGMILEMGLVDPKVRRLSALQHAASLAGIMLTCDMAELPEDKAMP-DMGMG	539
QY	541	MGMCM	544
Db	540	mgm	543
RESULT 9			
ID	AA75745		
XX	AA75745	standard; Protein; 544 AA.	
AC	AA75745:		
XX	21-MAR-2000	(first entry)	
DT	Neisseria gonorrhoeae ORF 982	protein sequence SEQ ID NO:2962.	
XX	Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;		
KW	antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;		
KW	antibacterial; gene therapy.		
OS	Neisseria gonorrhoeae.		
XX	MO9957280-AA2.		
XX	11-NOV-1999.		
PD	30-APR-1999;	99MO-US09346.	
..	01-MAY-1998;	98US-0083758.	
PR	31-JUL-1998;	98US-0094869.	
PR	02-SEP-1998;	98US-0098994.	
PR	02-SEP-1998;	98US-0099062.	
PR	09-OCT-1998;	98US-0103749.	
PR	09-OCT-1998;	98US-0103794.	
PR	09-OCT-1998;	98US-0103796.	
XX	25-FEB-1999;	99US-0121528.	
PA	(CHIR)	CHIRON CORP.	
PA	(GENO-)	INST GENOMIC RES.	
XX	Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M,		
PI	Petersen J, Plaza M, Rappelli R, Ratti G, Scalato E, Scarselli M,		
PI	Tellelin H, Venter JC;		
XX	WPI: 2000-062150/05.		
DR	N-PSDB: AA254507.		
XX	Novel Neisserial polypeptides predicted to be useful antigens for		
PT	vaccines and diagnostics		
XX	Claim 2; Page 1385; 1453pp; English.		
CC	AA253015 to AA254516, AA254517 to AA254615, and AA774253 to AA75941		
CC	represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides		

CC	and polypeptides AA554537 to AA254576 and AA254616 to AA25473 represent PCR primers used in the exemplification of the present invention.
CC	The polynucleotides, antibodies and compositions of CC
CC	polypeptides, the polynucleotides, antibodies and compositions of CC
CC	the invention can be used as vaccines, as diagnostic reagents, and as immunogenic compositions. The polypeptides can be used in the manufacture of medicaments for treating or preventing infection due to Neisseria bacteria (e.g. meningitis and septicemia), to detect the presence of Neisseria bacteria, or to raise antibodies. They may also have been used to screen for agonists or antagonists, which may themselves be used as antibacterial agents. The polynucleotides of the invention may also be used in gene therapy protocols.
SC	
XX	Sequence 544 AA:
Query Match	64.6%; Score 1750.5; DB 21; Length 544;
Best Local Similarity	64.9%; Pred. No. 1.5e-107;
Matches 353;	Conservative 78; Mismatches 112; Indels 1; Gaps 1
OY	1 MASKETLFDAKAREKLSRGVDKANAVKYTLGPKGRNVLEKSFGSVITRKDVSAAKET 60 :: :: :: Db 1 fassgmfrdnrlfkmvngvnlllpaaadvalagkgrnvvdrfgphltldkgvtvakel 60 OY 61 ELEDKFENNGAQWKKVKAARKTSDIADGGTTATVLAQATIREGVKKLVAAARNPMATRGI 120 :: Db 61 eltkdkfemmgagmrvkvaasktndvadgdltaclavlaqslvaeamkyvlaegmpldtkrgi 120 OY 121 DAAVAATVKTEEDSLTPPTPDOKELIAOVGTJANSDDTMIGMIIIEAAMAKVKKGVIYEBA 180 :: Db 121 dfaavaaveelknialkpdcstskeladygsfsasdevgalaiieamekvkgelylvedg 180 OY 181 KGLETTLDVVEGMKFEKDRGYLSPFYVTNPEKMGCELDNPYLICNEKITTSKMDLPILVOY 240 :: Db 181 ksleneidvvegmqrfdrylsyfyfindaeekqlagdlnpfvlildkkksinrldllpvleg 240 OY 241 AKVNPRLLIIADEEESALATLYNNKLRLGLQVVAYAKAPFEGERRRAMLEDIALITGGEA 300 :: Db 241 akasrpillilaeevgeaalatlvmnlrtglktvawkapgfgrtkamlqdialltggvv 300 OY 301 IFEDRCIKLENYSLSLGTAARVVIDKENNTIVDGAGKSSBDIKARVQOIRAOJEETSDDY 360 :: :: :: :: Db 301 lseevglsteaktlldlgqtktrejsentventvdgfydaeqleayvaeitrqletatsdy 360 OY 361 DREKIOLERLAKTVGSAVHVHGATETEEMKEKKRDVEDALINATRAAVEGIVPGCGAFV 420 :: Db 361 dkexikervaklagyaavilkvgatevemkekkrdevdalhatraevegvaggysvall 420 OY 421 RSTKVVLDDTKPADDDLACLNITRBSLEPLRIIANAGEGSIVEKVAEPKDGFQFNNA 480 :: :: :: :: Db 421 rtraalenihltgnaddagvqjlvlravesjlrtqvanaagepsvvnkvilegfygnyyna 480 OY 481 AGSEYEDLKAGVIDDKKTYTRIALNOANAASALLTTECIAAKPBPCKMPMPGGCMGC 540 :: :: :: :: Db 481 gsgseygdmlgmgylopakvrtselqnhaasalgaimitcdmaiaeipkecpavp-dimgmg 540 OY 541 MGGM 544 :: Db 540 mggm 543 RESULT 10 AAB19080 ID AAB19080 standard; protein; 545 AA. AAB19080; AAAB19080; 08-FEB-2001 (first entry) Amino acid sequence of a 60 kDa protein from Campylobacter jejuni. Antigenic protein; flagellaless mutant; vaccine; poultry; Campylobacter. Campylobacter jejuni.


```

OY 181 KGLSTLDVVEGKMFDRGYLSPYFVTNPEKMWCELDNPYLKNEKITSMKMDLPILQY 240
    ||: |||||:|||||||: || ||: || |||||: |||||: |||||: ||
DB 181 tglqdelvvegmqfdrgylspylfinkpetgavelspfllladkkslnremplveav 240
OY 241 AKVNRPLLIADVEGEALATLVNKLRLGALQVAVAKPFGERRKAMLEDIALITGGEA 300
    ||: |||||:|||||||: || ||: |||||: |||||: |||||: |||||: ||
DB 241 akagkplliadedvegalatlvntmrjlvkvaavkapfgdrkamladialitgvtv 300
OY 301 IFEDRGKILENVSLSIGTAKRVVIDKENTTYDGACKSEDIAKRVQIRAOIERTSSDY 360
    ||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: ||
DB 301 lseeigmelekatedlgygkrrvlnkdtllldgygeaalqgrvafqgrleaaady 360
OY 361 DREKLORLAKLVGVAVIHVGAATEEMKEKKDRVEDALNAPRAAEEGIVPGGTAFAV 420
    |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: ||
DB 361 dreklgervaklggvaivkgaaevemkekarrvedalnatraaeevvaaggyall 420
    |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: ||
OY 421 RSTKVLDDIKPADDELALGNIIRSLLEPLROIAANAGYEGSIIVEKREPPKDGFGFNA 480
    ||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: ||
DB 421 rvaskladlrgnedqvgjklvalrameaplrqvlncgeepsvantlvkgdgngyna 480
OY 481 ASGEYEDLIKAGVIDPKKVRIRALQNAASVSLLTTECAIAEKPEPKKMPMPG--GGM 538
    ||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: ||
DB 481 aleeegmldmgjldpklvrrsalqyaasvagimlttecmvdl--pkndaadlgaagm 538
OY 539 GGMGMDGM 547
    |||||: ||
DB 539 ggmngmgm 547

RESULT 12
ID AAM95001 standard; Protein: 548 AA.
XX
AC AAM95001;
XX
DT 12-MAY-1999 (first entry)
XX
DE MO9902989 Seq ID 10.
XX
KM Chapterone; molecular chaperone; groEL gene.
XX
OS Escherichia coli.
XX
MO9902989-A1.
XX
21-JAN-1999.
XX
PF 10-JUL-1998; 98MO-GB02042.
XX
PR 28-AUG-1997; 97GB-0018261.
XX
PR 10-JUL-1997; 97GB-0014582.
XX
PA (MED1-) MEDICAL RES COUNCIL.
XX
PI Chalellier J, Fersht A;
XX
DR WPI: 1999-121113/10.
XX
PT In vivo assay for the detection of chaperone fragment activity -
XX
PS useful for, e.g. complementing a molecular chaperone defect in vivo
XX
PS Disclosure: page 52-55; 64pp: English.
XX
CC The invention provides a method of providing chaperone activity in vivo,
XX
CC by administering to a cell a fragment of a molecular chaperone which has
XX
CC in vivo activity. Also provided is a method of determining whether a
XX
CC fragment of a molecular chaperone is active in vivo by: (i) providing a
XX
CC cell with a deficient molecular chaperone activity; (ii) administering
XX
CC the fragment to the cell; and (iii) determining whether the fragment
XX
CC complements the deficient endogenous molecular chaperone. The method can
XX
CC determine which fragments of the molecular chaperone can be used to
XX
CC complement a mutant or deficient molecular chaperone in vivo. The method

```

```

CC allows an in vivo assessment of chaperone activity as compared to
CC previous tests which only measured activity in vitro.
XX
SQ Sequence 548 AA;
Query Match 64.5%; Score 1746; DB 20; Length 548;
Best Local Similarity 64.5%; Pred. No. 3e-107;
Matches 354; Conservative 74; Mismatches 117; Indels 4; Gaps 2;
OY 1 MASKEILFDKAREKLSRGYDKLANAVKVTLPGRWVIEKSPGSPVITKDGYSVAKKEI 60
    ||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: ||
DB 1 maekdvkfgndarvkmjrgynvldavkvelgpggrvnlvdkstgaplckdgsvarel 60
OY 61 ELBDFENMKAQWYKVEAPRTSDIAGCTTATYLAQAIRREGYKLVAGRNPAIRGI 120
    |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: ||
DB 61 elcdkitemngqmwkvaskandaagqdtlatvlaqalllleglkavaagmpndlkrgj 120
OY 121 DKAAVAVTKELSDITKTRTPOKEIAQVGTISANDTTIGNIIEAMAKKVGCVITVEEA 180
    |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: ||
DB 121 dkavtaaveelkalsvpcsdskalaqvgctlsansdevgklllaeamkvkgelylvedg 180
OY 181 KGLSTLDVVEGKMFDRGYLSPYFVTNPEKMWCELDNPYLKNEKITSMKMDLPILQY 240
    ||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: ||
DB 181 tglqdelvvegmqfdrgylspylfinkpetgavelspfllladkkslnremplveav 240
OY 241 AKVNRPLLIADVEGEALATLVNKLRLGALQVAVAKPFGERRKAMLEDIALITGGEA 300
    ||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: ||
DB 241 akagkplliadedvegalatlvntmrjlvkvaavkapfgdrkamladialitgvtv 300
OY 301 IFEDRGKILENVSLSIGTAKRVVIDKENTTYDGACKSEDIAKRVQIRAOIERTSSDY 360
    ||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: ||
DB 301 lseeigmelekatedlgygkrrvlnkdtllldgygeaalqgrvafqgrleaaady 360
OY 361 DREKLORLAKLVGVAVIHVGAATEEMKEKKDRVEDALNAPRAAEEGIVPGGTAFAV 420
    |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: ||
DB 361 dreklgervaklggvaivkgaaevemkekarrvedalnatraaeevvaaggyall 420
OY 421 RSTKVLDDIKPADDELALGNIIRSLLEPLROIAANAGYEGSIIVEKREPPKDGFGFNA 480
    ||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: ||
DB 421 rvaskladlrgnedqvgjklvalrameaplrqvlncgeepsvantlvkgdgngyna 480
OY 481 ASGEYEDLIKAGVIDPKKVRIRALQNAASVSLLTTECAIAEKPEPKKMPMPG--GGM 538
    ||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: ||
DB 481 aleeegmldmgjldpklvrrsalqyaasvagimlttecmvdl--pkndaadlgaagm 538
OY 539 GGMGMDGM 547
    |||||: ||
DB 539 ggmngmgm 547

RESULT 13
ID AAB50536 standard; Protein: 548 AA.
XX
AC AAB50536;
XX
DT 16-MAR-2001 (first entry)
XX
DE Escherichia coli GroEL protein sequence SEQ ID NO:1.
XX
KM Escherichia coli; GroEL; mutant; muten; Improved stability;
XX
KM Chaperone; neurotropic; neuroprotective; Alzheimer's disease;
XX
KM prion disease; Creutzfeldt-Jacob's disease; CJD.
XX
OS Escherichia coli.
XX
PN MO200073463-A1.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000MO-GB02019.
XX

```


Query Match	Similarity	64.4%	Score 1743.5	DB 20	Length 544
Best Local	Similarity	62.9%	Pred. No. 4.4e-107		
Matches	346	Conservative	93	Mismatches	96
				Indels	13
				Gaps	4
QY	3	SKELFDKAREKRSRGVDKLANAVKVTLPGRNVVIEKSGSPVITKDGVSARETEL	62		
Db	2	akeklfseearrramrygdaladavkvcylgpkgrnrvlekkgsplltndygtlaketel	61		
QY	63	EDKFEENMGANQWKEYAPRTSDIAGDGTATVLAQAIYREGVKVLYAAGRNPAAIRGIDK	122		
Db	62	edafenmgaklvaaasklndvagggttatacltaeqamiregkknvtaagapvyrkmgneq	121		
QY	123	AVNAVTKELSLITPKTRROCKELAQVCTISANSPTTIGNIIEAMAKVKCGVITVEAKG	182		
Db	122	avavaieiikelstskprieqkeslaqvaaista-adeevgsilaeamervgndyvlileesky	180		
	183	LETTLDVVEGKMKFORGYLSPIYFTNPKEKMCLEDPYLLCNEKRTTSMKMLPILEQYAK	242		
	181	ftcteevegmgfdtrgyaspyrmtsdskmeavldnpylltdckltlntgelipvtieqvq	240		
QY	243	VNRPLLIATEDVEGEGALTTLVVNNKIRGALQVAVAKAPFGERRKAMLEDIALTSGEALF	302		
Db	241	gqkprllilaedvegeaalalvnnkrlrgfnavaavkapfgdrkxamledlavltgevalt	300		
QY	303	EDRCKILENLSLSLTGATKRRVYIDKENTTYIDGACKSEDIKARVQIAQIIEFTSSDIDR	362		
Db	301	edlgltdlkslqagrkaskvvvteentllvegagetklasrvlqlaqveetsefldr	360		
QY	363	EKLDERLAKLVGGVAVIVGATFEREMKEKDRYDADLANTPAAVEGIVPGCGTAFRS	422		
Db	361	eklgerlrlaklgyvavvlkvgaatecelrklrtlledalinsttraaavegivsrggtalvny	420		
QY	423	IKVLDDIKPADDDDELAGLNIIRSLSEEPRLIAANAGYEGSVIYKVEPRPKDGFGNAS	482		
Db	421	ynkvaave-aegdaqgtlnlvrlraleerplrqhmaglegsvlverlkneelgvgnaat	479		
QY	483	GEYEDLRKAGVYIDKKVYRIALQNAASVASLLTTTECAIAEKPEPKKMPMGCG-----	537		
Db	480	gewvmltekqivdpklvtrlsalqnaasvaamlltleavvadkpeen-----gggaqmpd	533		
QY	538	MGGMGMDGM 547			
Db	534	mgmgmgmgmg 543			
	ULT 15				
	A61378				
	AAW61378 standard; Protein: 548 AA.				
XX	AAW61378:				
XX	AC				
XX	AD				
XX	AE				
XX	AF				
XX	AG				
XX	AH				
XX	AI				
XX	AK				
XX	AL				
XX	AM				
XX	AN				
XX	AO				
XX	AP				
XX	AQ				
XX	AR				
XX	AS				
XX	AT				
XX	AAW61378:				
XX	AC				
XX	AD				
XX	AE				
XX	AF				
XX	AG				
XX	AH				
XX	AI				
XX	AK				
XX	AL				
XX	AM				
XX	AN				
XX	AO				
XX	AP				
XX	AQ				
XX	AR				
XX	AS				
XX	AT				
XX	AAW61378:				
XX	AC				
XX	AD				
XX	AE				
XX	AF				
XX	AG				
XX	AH				
XX	AI				
XX	AK				
XX	AL				
XX	AM				
XX	AN				
XX	AO				
XX	AP				
XX	AQ				
XX	AR				

xx New chapterone polypeptides - comprising fragments of a GroEL
 pt sequence, used for refolding, reactivating or reconditioning
 pt proteins, e.g. for treating Alzheimer's disease
 xx
 xx Claim 1; Fig 7; 97pt; English.
 xx
 CC GroEL is a member of the hsp60 heat shock protein family which
 CC facilitates the folding of proteins by preventing aggregation and by
 CC continuously annealing misfolded proteins by unfolding them to a state
 CC from which refolding can start again. The GroEL protein sequence is used
 CC to create chaperone polypeptides which have the ability to refold,
 CC reactivate or recondition proteins. The chaperone polypeptide products
 CC can be used for the treatment of diseases associated with aberrant
 CC protein/polypeptide structure such as Alzheimer's disease.
 xx
 SQ Sequence 548 AA:
 xx
 Query Match 64.3%; Score 1741; DB 19; Length 548;
 Best Local Similarity 64.3%; Pred. No. 6.4e-107;
 Matches 353; Conservative 74; Mismatches 116; Indels 4; Gaps 2.

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-467-822-31

Query Match 68.1%; Score 1844; DB 2; Length 548;
Best Local Similarity 67.2%; Pred. No. 1.1e-153;
Matches 366; Conservative 81; Mismatches 96; Indels 2; Gaps 1;

QY 3 SKELLFPAKREKLSRGVDLANAVKVTLCPKGRNVIEKSPSPVTKGVSVAKEIEL 62
DB 2 AKELRFDDARLQMLAGVNLADAVQVTMPGRGNVLEKSYGAPVTYTKGVSVAKEIEF 61
QY 63 EDFENKGAOMVEVAPKTSIDAGDGTATVLAQAATYREGVKLVAGRNPMIAIKRGIDK 122
DB 62 EHRFMNNGAOMVEVASKTSDTAGDGTATVLAARSILVEGHKVAAGRNPMIDKRGIDK 121
QY 123 AVVAVTKEISDITKPTRDQKEIAOVGTISANSOTTIGNIIAEAMAKVKGCVITVEEAKG 182
DB 122 AVLAVTKRLQAMSKPCDSKAIAOVGTISANSDEAIGAITAEAMEKVGKGVITVEEAKG 181
183 LETTLDVVEGMKFDGRGLSPYFVTNPENKVCEDLNPYILCNEKKITSMKMLPILVEOVAK 242
182 LENELVVEGMQFDGRGIIISPYFINNOOMSCLEHPIILLVDKVVSSIREMLSVLEGVAK 241
QY 243 VNRPLIIAEDVEGEALATLVNKLRGALOVAVKAPGFERRKAMLEDIAILTGEAIF 302
DB 242 SGRPLIIAEDIEGEALATLVNMMRGIVKCAVKAPGFDRRKAMLODIAILTKGVIS 301
QY 303 EDRGIKLENVSLSLGTAKRNVIDKENTTIVDAGKSEDIKARVKQIRAOIEETSSDYDR 362
DB 302 EEIKSLLEGATLEDGSAKRIIVTKENTTIIIDGKATEINARIAQIRAOIEETSSDYDR 361
QY 363 EKLQERLAKLVGVAIVHGAATEETEMKEKKDRVEDALNATRAVEGIIYPGGGTAVRS 422
DB 362 EKLQERLAKLVGVAIVHGAATEVEEMKEKKARVEDALNATRAVEGIIYAGGVALIRA 421
QY 423 IKVLDIKPADDELACINIIIRSLERPLQOIAANAGYEGSIVVEKREKPDGFGFNAAS 482
DB 422 OKALDSLKGNDQDMGMINILRAIESPMQIYVNAEYVAVKAEHNDNGFNAAT 481
QY 483 GEYEDLIKAGYIDPKKVTIRALONAAVASLILTECAIAEKPEPKKDMPPGGMGMG 542
DB 482 GEYGDVEMGILDPKVTIRALONAAVASLILTECMVADL--PKKEGVAGADMGMG 539
QY 543 GMDGM 547
DB 540 GMDGM 544

11LT 2
38-432-697-31
Sequence 31, Application US/08432697
Patent No. 6248330
GENERAL INFORMATION:
APPLICANT: Labigne, Agnes
APPLICANT: Sauerbaum, Sebastien
APPLICANT: Ferrero, Richard L.
APPLICANT: Thiberge, Jean-Michel
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
TITLE OF INVENTION: HELICOBLACTER INFECTION, POLYPEPTIDES FOR USE IN THE
TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08-432,697
FILING DATE: 02-MAY-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495.0137-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-432-697-31

Query Match 68.1%; Score 1844; DB 4; Length 548;
Best Local Similarity 67.2%; Pred. No. 1.1e-153;
Matches 366; Conservative 81; Mismatches 96; Indels 2; Gaps 1;

QY 3 SKELLFPAKREKLSRGVDLANAVKVTLCPKGRNVIEKSPSPVTKGVSVAKEIEL 62
DB 2 AKELRFDDARLQMLAGVNLADAVQVTMPGRGNVLEKSYGAPVTYTKGVSVAKEIEF 61
QY 63 EDFENKGAOMVEVAPKTSIDAGDGTATVLAQAATYREGVKLVAGRNPMIAIKRGIDK 122
DB 62 EHRFMNNGAOMVEVASKTSDTAGDGTATVLAARSILVEGHKVAAGRNPMIDKRGIDK 121
QY 123 AVVAVTKEISDITKPTRDQKEIAOVGTISANSOTTIGNIIAEAMAKVKGCVITVEEAKG 182
DB 122 AVLAVTKRLQAMSKPCDSKAIAOVGTISANSDEAIGAITAEAMEKVGKGVITVEEAKG 181
QY 183 LETTLDVVEGMKFDGRGLSPYFVTNPENKVCEDLNPYILCNEKKITSMKMLPILVEOVAK 242
DB 182 LENELVVEGMQFDGRGIIISPYFINNOOMSCLEHPIILLVDKVVSSIREMLSVLEGVAK 241
QY 243 VNRPLIIAEDVEGEALATLVNKLRGALOVAVKAPGFERRKAMLEDIAILTGEAIF 302
DB 242 SGRPLIIAEDIEGEALATLVNMMRGIVKCAVKAPGFDRRKAMLODIAILTKGVIS 301
QY 303 EDRGIKLENVSLSLGTAKRNVIDKENTTIVDAGKSEDIKARVKQIRAOIEETSSDYDR 362
DB 302 EEIKSLLEGATLEDGSAKRIIVTKENTTIIIDGKATEINARIAQIRAOIEETSSDYDR 361
QY 363 EKLQERLAKLVGVAIVHGAATEETEMKEKKDRVEDALNATRAVEGIIYPGGGTAVRS 422
DB 362 EKLQERLAKLVGVAIVHGAATEVEEMKEKKARVEDALNATRAVEGIIYAGGVALIRA 421
QY 423 IKVLDIKPADDELACINIIIRSLERPLQOIAANAGYEGSIVVEKREKPDGFGFNAAS 482
DB 422 OKALDSLKGNDQDMGMINILRAIESPMQIYVNAEYVAVKAEHNDNGFNAAT 481
QY 483 GEYEDLIKAGYIDPKKVTIRALONAAVASLILTECAIAEKPEPKKDMPPGGMGMG 542
DB 482 GEYGDVEMGILDPKVTIRALONAAVASLILTECMVADL--PKKEGVAGADMGMG 539
QY 543 GMDGM 547
DB 540 GMDGM 544

RESULT 3
US-08-467-822-32

QY 301 IFEDRGIKLENSLSIGTAKRVYIDKENTTIYDGAACKSEDIKARVQIRAOIEETSSDY 360
Db 301 ISEIGMELEKATLEDLGCAKRVYINKOTTTIIDVGEBAIOGRAVOIRQOIEEATSDY 360
QY 361 DREKLOERLAKLVGVAIVHGAATETEMKEKKRDVEDALNATRAAVEGIVPGGTAFF 420
Db 361 DREKLOERVAKLAGVAIVHGAATEVEEMKEKKARVEDALHATRAAVEGIVAGGVALLI 420
QY 421 RSIVKIDDIKPADDDDELACINIRSLIEEPLRQIAANAGYEGSIYVEKREPKDGFQFNA 480
Db 421 RVASKIADLKGONEDONVGIKVALRAMEAPLRQIYLVNCGEPPSVANTVYGGDNGYNA 480
QY 481 ASGEEDLIKAGVIDPKKVTIRIALQNAASVASILLTTECAIAEKPEPKDMPMG--GGM 538
Db 481 ATEEXGNIMDGLDPTTKVTRSALOYAASVAGLMTTECMVTDL--PKNDADLGAAGM 538
QY 539 GGMGMDGM 547
Db 539 GGMGMDGM 547

RESULT 5

us-08-432-697-32
Sequence 32, Application US/08432697

Patent No. 6248330

GENERAL INFORMATION:

APPLICANT: Labigne, Agnes

APPLICANT: Sauetbaum, Sebastien

APPLICANT: Fierro, Richard L.

APPLICANT: Thibierge, Jean-Michel

TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST

TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE

TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID

TITLE OF INVENTION: POLYPEPTIDES

NUMBER OF SEQUENCES: 44

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

ADDRESSEE: Dunner

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PC-DOS/MS-DOS

CURRENT APPLICATION DATA: Patent Release #1.0, Version #1.30

APPLICATION NUMBER: US/08/432.697

FILING DATE: 02-MAY-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Meyers, Kenneth J.

REGISTRATION NUMBER: 25,146

REFERENCE/DOCKET NUMBER: 03495.0137-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 408-4000

TELEFAX: (202) 408-4400

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 548 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-432-697-32

Query Match 64.3%; Score 1741; DB 4; Length 548;
Best Local Similarity 64.3%; Pred. NO. 1.3e-144;
Matches 353; Conservative 74; Mismatches 118; Indels 4; Gaps 2;

QY 1 MASKEILFDKAREKLSRGVDKLANAVKTVLGRGRNVIEKSGSPVITKDGVSVAKEI 60
Db 1 MAADVKFEGNDARKVKRNVNLADAVKVLGPKGRNVVLDKSGFAPITTKDGVSAKEI 60
QY 61 ELDKFEENMGAAQVKEVAPKPTSDIAGCGTTATVLAQAIYRECVKLVAAQRNMAIKRGI 120
Db 61 ELDKFEENMGAAQVKEVAPKPTSDIAGCGTTATVLAQAIYRECVKLVAAQRNMAIKRGI 120
QY 121 DKAIVAVTKELSDITKPTPDQKEIAQVGTISANSPTTIGNIAEAMAKVKGKGVITVEEA 180
Db 121 DKAIVAVTKELSDITKPTPDQKEIAQVGTISANSPTTIGNIAEAMAKVKGKGVITVEEA 180
QY 181 KGLTTLVDVEGKKEFDKRGVTSPIFTVNPKEKMGELDNPYLLCNKRTSKMDLPTLEQV 240
Db 181 TGLDDELVDVEGKKEFDKRGVTSPIFTVNPKEKMGELDNPYLLCNKRTSKMDLPTLEQV 240
QY 241 AKVNRPLIIAEDVEGEALATLVNKLRLGALQVAAKAGFGRRRAMLEDIAIITLGEA 300
Db 241 AKVNRPLIIAEDVEGEALATLVNKLRLGALQVAAKAGFGRRRAMLEDIAIITLGEA 300
QY 301 IFEDRGIKLENSLSIGTAKRVYIDKENTTIYDGAACKSEDIKARVQIRAOIEETSSDY 360
Db 301 ISEIGMELEKATLEDLGCAKRVYINKOTTTIIDVGEBAIOGRAVOIRQOIEEATSDY 360
QY 361 DREKLOERLAKLVGVAIVHGAATETEMKEKKRDVEDALNATRAAVEGIVPGGTAFF 420
Db 361 DREKLOERVAKLAGVAIVHGAATEVEEMKEKKARVEDALHATRAAVEGIVAGGVALLI 420
QY 421 RSIVKIDDIKPADDDDELACINIRSLIEEPLRQIAANAGYEGSIYVEKREPKDGFQFNA 480
Db 421 RVASKIADLKGONEDONVGIKVALRAMEAPLRQIYLVNCGEPPSVANTVYGGDNGYNA 480
QY 481 ASGEEDLIKAGVIDPKKVTIRIALQNAASVASILLTTECAIAEKPEPKDMPMG--GGM 538
Db 481 ATEEXGNIMDGLDPTTKVTRSALOYAASVAGLMTTECMVTDL--PKNDADLGAAGM 538
QY 539 GGMGMDGM 547
Db 539 GGMGMDGM 547

RESULT 6

us-08-470-260-6

Sequence 6, Application US/08470260

Patent No. 6077706

GENERAL INFORMATION:

APPLICANT: Covacci, Antonello

APPLICANT: Bugnoli, Massimo

APPLICANT: Telford, John

APPLICANT: Macchia, Giovanni

APPLICANT: Rappuoli, Rino

TITLE OF INVENTION: Helicobacter Pylori Proteins Useful

TITLE OF INVENTION: for Vaccines and Diagnostics

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Chiron Corporation

STREET: 4560 Horton Street

CITY: Emeryville

STATE: California

COUNTRY: USA

ZIP: 94608-2916

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PC-DOS/MS-DOS

CURRENT APPLICATION DATA: Patent Release #1.0, Version #1.25

APPLICATION NUMBER: US/08/470.260

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/256,848

FILING DATE: 21-OCT-1994

: ATTORNEY/AGENT INFORMATION:
: NAME: McClung, Barbara G.
: REGISTRATION NUMBER: 33,113
: REFERENCE/DOCKET NUMBER: 0316.001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (510) 601-2708
: TELEFAX: (510) 655-3542
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 546 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-470-260-6

Query Match	63.6%	Score	1723.5	DB	3	Length	546
Best Local Similarity	64.1%	Pred. No.	4.3e-143				
Matches	350	Conservative	73	Mismatches	120	Indels	3
						Gaps	2

[illegible]

```

RESULT 7
US-08-471-491-6
: Sequence 6, Application US/08471491B
: Patent No. 6090611
: GENERAL INFORMATION:
: APPLICANT: Covacci, Antonello
: APPLICANT: Telford, John
: APPLICANT: Macchia, Giovanni
: APPLICANT: Rappuoli, Rino
: TITLE OF INVENTION: Helicobacter pylori Proteins Useful For Vaccines And

```

```

? TITLE OF INVENTION: Diagnostics
? FILE REFERENCE: CHIR004
? CURRENT APPLICATION NUMBER: US/08/471,491B
? CURRENT FILING DATE: 1995-06-06
? NUMBER OF SEQ ID NOS: 8
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 6
? LENGTH: 546
? TYPE: PR1
? ORGANISM: Helicobacter pylori
US-08-471-491-6

```

Query Match	63.68%	Score 1723.5	DB 3	Length 546
Best Local Similarity	64.1%	Pred. No. 4.3e-143		
Matches 350; Conservative	73;	Mismatches 120;	Indels 3;	Gaps 2

[illegible]

RESULT : 8
 US-08-466-662-6
 ; Sequence 6, Application US/08466662B
 ; Patent No. 6130059
 ; GENERAL INFORMATION:
 ; APPLICANT: Covacci, Antonello
 ; APPLICANT: Bugnoli, Massimo
 ; APPLICANT: Telford, John
 ; APPLICANT: Macchia, Giovanni
 ; APPLICANT: Rappuoli, Rino
 ; TITLE OF INVENTION: Helicobacter Pylori Proteins Useful For Vaccines And
 ; TITLE OF INVENTION: Diagnostics
 ; FILE REFERENCE: CHIR0057
 ; CURRENT APPLICATION NUMBER: US/08/466,662B
 ; CURRENT FILING DATE: 1995-06-06
 ;

Matches 352: Conservative 74; Mismatches 115; Indels 11; Gaps 4;

QY 1 MASSEILFDKAKAREKLSRGVDKLANAVKVTLGPKGRNVIEKSGSPYITKDGVSVAKEI 60
 Db 1 MAADVDFGNDARVKMGRGVNLDAVKVTLGPKGRNVLDKSGFAPITIKDGVSAKEI 60

QY 61 ELLEKPFENMGAGMVEVAPRTSDIAGDGTATVLAQAIRESKVLVAAGRNPAIRGI 120
 Db 61 EPEDKFEKMGAGMVEVAPRTSDIAGDGTATVLAQAIRESKVLVAAGRNPAIRGI 120

QY 121 DKAVAATKELSDITKPTPDQKEIAQVGTISANSDDTIGNIIAEAMAKVGKGYITVEEA 480
 Db 121 DKAVAATKELSDITKPTPDQKEIAQVGTISANSDDTIGNIIAEAMAKVGKGYITVEEDG 180

QY 181 KGLTTLDVVEGKMFEDRGVLSPEFVTNPEKRVCELDNPYLCEKKTSMKMDLPILQV 240
 Db 181 TGLDELDVVEGKMFEDRGVLSPEFVTNPEKRVCELDNPYLCEKKTSMKMDLPILQV 240

QY 241 AKNRPILLIIEVDEGALTLVNNKRGALQVAVAPGFERKRMALDIALITGGEA 300
 Db 241 AKNRPILLIIEVDEGALTLVNNKRGALQVAVAPGFERKRMALDIALITGGEA 300

QY 301 IFEDRGKILENVSLSGTAKRVVIDKENTTIVDAGKSEDIKARVQIAQIETSSDY 360
 Db 301 ISEIGMELEKATLEDGQAKRVINKDTTIIIDGCEALIOGRVAQIQIIEATSDY 360

QY 361 DREKLOERLAKLVGAIVHGAATETEMKEKKDRVEDALNATRAVEEGIVGGGAFAV 420
 Db 361 DREKLOERLAKLVGAIVHGAATETEMKEKKDRVEDALNATRAVEEGIVGGGVALI 420

QY 421 RSJIKVLDIIRKPADDELACINIR---RSLEPLRQIAAAGYEGSIYKVEKPEKDFG 477
 Db 421 RVASKIADLQGNEDQ---NVSSSLRAMEAPLRQIVLNGCEPSSVANTVAVGGGNGY 476

QY 478 FNAASEYEDLIIAGYIDPKRTIRIALONASVASLLITTECAIAEKPEPKMPMG-- 535
 Db 477 YNATEEYCGMILMGJLIDPKRTIRIALONASVASLLITTECAIAEKPEPKMPMG-- 534

QY 536 GGMGGMGMDGM 547
 Db 535 GGMGGMGMDGM 546

RESULT 12
 US-08-467-822-30
 ; Sequence 30, Application US/08467822
 ; Patent No. 5843460
 ; GENERAL INFORMATION:
 ; APPLICANT: Labigne, Agnes
 ; APPLICANT: Sauetbaum, Sebastien
 ; APPLICANT: Ferrero, Richard L.
 ; APPLICANT: Thibierge, Jean-Michel
 ; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
 ; TITLE OF INVENTION: HELICOCHACER INFECTION, POLYPEPTIDES FOR USE IN THE
 ; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
 ; NUMBER OF SEQUENCES: 44
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
 ; ADDRESSEE: Dunner
 ; STREET: 1300 I Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3315
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/467,822
 ; FILING DATE: 06-JUN-1995

CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/447,177
 FILING DATE: 19-MAY-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/432,697
 FILING DATE: 02-MAY-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Meyers, Kenneth J.
 REGISTRATION NUMBER: 25,146
 REFERENCE/DOCKET NUMBER: 03495, 0137-02000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 408-4000
 TELEFAX: (202) 408-4400
 INFORMATION FOR SEQ. ID NO: 30:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 545 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: Protein
 LOCATION: 1..545
 OTHER INFORMATION: /product= "H. pylori - Hsp B."

US-08-467-822-30

Query Match 62.9%; Score 1702; DB 2; Length 545;
 Best Local Similarity 63.7%; Pred. No. 3,3e-141;
 Matches 348; Conservative 73; Mismatches 121; Indels 4; Gaps 3;

QY 3 SKRIILDARAREKLSGVDKLANAVVTLGPKGRNVIEKSGSPYITKDGVSVAKEI 62
 Db 2 AKKIKFSDSARNLIFEGVRLHDVAVVTWMPGRGNVLIOKSVGAPITKDGVSVAKEI 61

QY 63 EDKFNMGAGMVEVAPKTSIDNGDTTATVLAQAIRESKVLVAAGRNPAIRKGI 122
 Db 62 SCPVANMGAGMVEVAPKTSIDNGDTTATVLAQAIRESKVLVAAGRNPAIRKGI 121

QY 123 AVAVATKELSDITKPTPDQKEIAQVGTISANSDDTIGNIIAEAMAKVGKGYITVEEAG 182
 Db 122 APEAIINLEKASKKVGKEEITQVATISANSDDHNGIKLADMEKVGKGYITVEEAG 181

QY 183 LETTLVVEGKMFEDRGVLSPEFVTNPEKRVCELDNPYLCEKKTSMKMDLPILQV 242
 Db 182 IEDELVDVVEGKMFEDRGVLSPEFVTNPEKRVCELDNPYLCEKKTSMKMDLPILQV 241

QY 243 VNPFLIIEVDEGALTLVNNKRGALQVAVAPGFERKRMALDIALITGGEA 302
 Db 242 EGRPLIIEVDEGALTLVNNKRGALQVAVAPGFERKRMALDIALITGGEA 301

QY 303 EDRCIKLENVSLSGTAKRVVIDKENTTIVDAGKSEDIKARVQIAQIETSSDY 362
 Db 302 EELGLLENAVEEFLGKAK-IVIDKNTTIVDAGKSHDYKDRVAQIQIISTSDY 360

QY 363 EKIQERLAKLVGAIVHGAATETEMKEKKDRVEDALNATRAVEEGIVGGGAFAV 422
 Db 361 EKIQERLAKLVGAIVHGAATETEMKEKKDRVEDALNATRAVEEGIVGGGALIRA 420

QY 423 IKVLDDIKPADDELACINIR---RSLEPLRQIAAAGYEGSIYKVEKPEKDFG 482
 Db 421 AQKVH--LNLHDEKGYELIRAIRAPLAQAIINNGYGGVAVNEVEKHEGFGNASH 478

QY 483 GEYEDLIKAGVIDPKRTIRIALONASVASLLITTECAIAEKPEPKMPMG--GGMGM 541
 Db 479 GGYVDFKKEGIIIDPKRTIRIALONASVASLLITTEATYHEIKKEKAAPAMPDMGMDGM 538

QY 542 GGMGMDGM 547
 Db 539 GGMGMDGM 544

Matches 326; Conservative 83; Mismatches 130; Indels 4; Gaps 4;

```

OY 3 SKELFDKAREKLSRGVDKLANAVKVTLPKGRNVVIEKSPGSPVITKGVSAKEIEL 62
DB 2 AKTIADVEARRGLEKRLNLADAVKTLCPKGNVLEKWKWGPITTDGVSIAKEIEL 61
OY 63 EDFENMGOMVEVAPKTSIDAGDGTATVLAQAIYRGVKKVLAAGRPMIAIKRIDK 122
DB 62 EDYERKIGALVEKAKTIDVAGDGTATVLAQALVREGLRVAVAGANPLGLKRIEX 121
OY 123 AVAVATKELSDITKPYRDOKEIAOVGTISANSOTTCNIIAEAMAVKGVITVEBAK 182
DB 122 AVEAVTOSLKSKEVETKEQISATAISA-GDTQIGELIAEAMDKVNGCVITVEESNT 180
OY 183 LETTLDVVEGKMFDRGLSPYFVTNPENKMYCELDPYIILCNEKITSKMDLPLEQVAK 242
DB 181 FGIQLELTGEMRFDKGISGFVTDAREQAVLEDPYIILVSSKVSIVKDLPLEKVIQ 240
OY 243 VNRPLLIADVEGEALATLVNKLKGALOVAVKAPGEGERRKAMLEDAITLGGEAIF 302
DB 241 AGKPLLIADVEGEALSTLVNKKIRGTEKSVAVKAPGEGDRKKAMLOMAITLGGQVVS 300
OY 303 EDGKIKLENVSLSLGTAKRVVIDKENTTVI DGAGKSEDIKARVKQIRAOIETSSDYDR 362
DB 301 ERVGLSLETADVSLGQARKVYVTKDETTIVEGSGSDAIAGRVAOIRAEIENSDDYDR 360
OY 363 EKIOERLAKLVGVAVIHGAATEEMKEKKDRVEDALNATRAVEGIVPGGTAFAVRS 422
DB 361 EKIOERLAKLVGVAVIKGAATEVELEKERRHIEDAVRNKAKAVEEGIVAGGVALLOS 420
OY 423 IKVLDDIKPADDELAGLNIIRSLSEPLRQIAANAGYEGSIVVEYREPKDGFNMAAS 482
DB 421 APALDDL-GLTGDEATGANIVRALAPLKOIAFNGLEGCVAAEKVSNLPAGHCLMAAT 479
OY 483 GEYEDLIKAGVIDPKKVTIRALONASVASLLITTECAIAEKPEPKKMPGCGMGNG 542
DB 480 GEYEDLIKAGVADPKVTRSAIONASIALFLITTEAVVADKPE-KASAP-AGDPGGMG 537
OY 543 GMD 545
DB 538 GMD 540

```

RESULT 15
US-08-997-362-160
Sequence 160 Application US/08997362
Patent No. 5985287
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Miyama, Jun
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Scott, Linda
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
NUMBER OF SEQUENCES: 194
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA: US/08/997,362
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/8/3,970
FILING DATE: June 12, 1997
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/7/05,347
FILING DATE: August 29, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ. ID NO: 160:
SEQUENCE CHARACTERISTICS:
LENGTH: 541 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-997-362-160

Query Match 59.0%; Score 1597; DB 2; Length 541;
Best Local Similarity 60.0%; Pred. No. 5,4e-132.
Matches 326; Conservative 83; Mismatches 130; Indels 4; Gaps 4;

```

OY 3 SKELFDKAREKLSRGVDKLANAVKVTLPKGRNVVIEKSPGSPVITKGVSAKEIEL 62
DB 2 AKTIADVEARRGLEKRLNLADAVKTLCPKGNVLEKWKWGPITTDGVSIAKEIEL 61
OY 63 EDFENMGOMVEVAPKTSIDAGDGTATVLAQAIYRGVKKVLAAGRPMIAIKRIDK 122
DB 62 EDYERKIGALVEKAKTIDVAGDGTATVLAQALVREGLRVAVAGANPLGLKRIEX 121
OY 123 AVAVATKELSDITKPYRDOKEIAOVGTISANSOTTCNIIAEAMAVKGVITVEBAK 182
DB 122 AVEAVTOSLKSKEVETKEQISATAISA-GDTQIGELIAEAMDKVNGCVITVEESNT 180
OY 183 LETTLDVVEGKMFDRGLSPYFVTNPENKMYCELDPYIILCNEKITSKMDLPLEQVAK 242
DB 181 FGIQLELTGEMRFDKGISGFVTDAREQAVLEDPYIILVSSKVSIVKDLPLEKVIQ 240
OY 243 VNRPLLIADVEGEALATLVNKLKGALOVAVKAPGEGERRKAMLEDAITLGGEAIF 302
DB 241 AGKPLLIADVEGEALSTLVNKKIRGTEKSVAVKAPGEGDRKKAMLOMAITLGGQVVS 300
OY 303 EDGKIKLENVSLSLGTAKRVVIDKENTTVI DGAGKSEDIKARVKQIRAOIETSSDYDR 362
DB 301 ERVGLSLETADVSLGQARKVYVTKDETTIVEGSGSDAIAGRVAOIRAEIENSDDYDR 360
OY 363 EKIOERLAKLVGVAVIHGAATEEMKEKKDRVEDALNATRAVEGIVPGGTAFAVRS 422
DB 361 EKIOERLAKLVGVAVIKGAATEVELEKERRHIEDAVRNKAKAVEEGIVAGGVALLOS 420
OY 423 IKVLDDIKPADDELAGLNIIRSLSEPLRQIAANAGYEGSIVVEYREPKDGFNMAAS 482
DB 421 APALDDL-GLTGDEATGANIVRALAPLKOIAFNGLEGCVAAEKVSNLPAGHCLMAAT 479
OY 483 GEYEDLIKAGVIDPKKVTIRALONASVASLLITTECAIAEKPEPKKMPGCGMGNG 542
DB 480 GEYEDLIKAGVADPKVTRSAIONASIALFLITTEAVVADKPE-KASAP-AGDPGGMG 537
OY 543 GMD 545
DB 538 GMD 540

```

Search completed: August 1, 2001, 17:18:09
Job time: 2460 sec

Thu Aug 2 08:05:01 2001

us-09-077-574a-2.rai

Page 11

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2001, 16:48:44 ; Search time 40.91 Seconds

(without alignments)
1020.378 Million cell updates/sec

Title: US-09-077-574A-2

Perfect score: 2708

Sequence: 1 MASKELLFPAKAREKLSRCV.....KIDMPGCGMGCGMDGMV 548

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

-sarched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1866	68.9	551	2	JC2562
2	1863	68.8	545	2	JN0509
3	1862.5	68.8	547	2	F82783
4	1852.5	68.4	546	2	B47073
5	1845	68.1	550	2	A41468
6	1844.5	68.1	546	2	S65596
7	1843.5	68.1	547	2	I40331
8	1843.5	68.1	552	2	S39765
9	1828.5	67.5	544	2	S23918
10	1825.5	67.4	547	2	B83098
11	1823	67.3	546	2	S23347
12	1823	67.3	550	2	S35309
13	1822	67.3	546	2	I40342
14	1817.5	67.1	544	2	B36917
15	1816.5	67.1	541	2	JN0512
16	1814.5	67.0	544	2	H81964
17	1814.5	67.0	549	2	JC2564
18	1812.5	66.9	545	2	S47530
19	1810.5	66.9	544	2	S37039
20	1808.5	66.8	544	2	C81021
21	1805.5	66.7	547	2	B43606
22	1802.5	66.6	544	2	S61300
23	1799	66.4	546	2	B54539
24	1794.5	66.3	544	2	S61302
25	1794	66.2	546	2	S35311
26	1791	66.1	541	2	S61301
27	1789.5	66.1	541	2	JN0511
28	1788.5	66.0	544	2	S61303
29	1786	66.0	545	2	C70489

30	1780.5	65.7	544	2	JC6063	chaperonin groEL -
31	1780.5	65.7	544	2	B83720	class I heat-shock
32	1774.5	65.5	545	2	S51563	heat shock protein
33	1773	65.5	544	1	B43827	chaperonin groEL -
34	1765	65.2	546	2	S34938	heat shock protein
35	1761.5	65.0	544	2	B82048	chaperonin, 60 kD
36	1759	65.0	548	2	C64076	chaperonin groEL -
37	1755.5	64.8	539	2	B49855	heat shock protein
38	1755	64.8	545	2	G81328	60 kD chaperonin (
39	1754.5	64.8	543	2	JC5771	chaperonin groEL-1
40	1747	64.5	544	2	S70667	chaperonin groEL -
41	1746	64.5	548	1	BVBCGL	chaperonin groEL -
42	1745	64.4	547	2	JC4519	heat-shock protein
43	1745	64.4	548	2	B86110	hypothetical prote
44	1743.5	64.4	544	2	B41884	58k heat shock pro
45	1743	64.4	550	2	S26423	heat shock protein

ALIGNMENTS

RESULT 1

JC2562 chaperonin groEL protein - Amoeba proteus

C:Species: Amoeba proteus

C>Date: 17-May-1995 #sequence_revision 14-Jul-1995 #text_change 26-Aug-1999

C:Accession: JC2562

R:Ahm, T.I.: Lim, S.T.: Leeu, H.K.: Lee, J.E.: Jeon, K.W.

Gene 14811281, 43-49, 1994

A:Title: A novel strong promoter of the groEL operon of symbiotic bacteria in Amoe

A:Reference number: JC2561

A>Note: due to a typographical error the volume number 148 appears as 128

A:Accession: JC2562

A:Molecule type: DNA

A:Residues: 1-551 <AHN>

A:Cross-References: GB:M86549; NID:9155400; PIDN:AC09381.1; PID:9155402

C:Comment: This protein is involved in the assembly of oligomeric protein complexes

C:Genetics:

A:Gene: groELX

C:Superfamily: chaperonin groEL

C:Keywords: molecular chaperone

Query Match	Score	Length	DB 2:
Best Local Similarity	67.9%	Pred. No. 166-83;	
Matches 373; Conservative 74; Mismatches 96; Indels 6; Gaps 2;			
QY 3	SKELLFPAKAREKLSRGVDKLANAVKVTGPKGRNVIEKSPVITRDGVSAKEIETL 62		
DB 2	AKELRGDARQOMLAGVNAADRVKATGPGSRNVILERSFGAPVITRDGVSAKEIEF 61		
QY 63	EDKFNMGAMQVEVAPKTSIAGDGTATATVLAQIYREGVKLVAGRNPAIKRGIDK 122		
DB 62	ENFRKMGAMQVEVAKTSIDFGDGTATATVLAIRSVIEGHKVAAGMHPDLKRGIDK 121		
QY 123	AVAVATKELSDITKPRROCKELTAOVGTISANSPTTGINIAEAMVAGKGVITVEBAG 182		
DB 122	AVATATKELQKSKPKCGALNAOVGTISANSQALGSIIEAMKVEVAGKGVITVEBAG 181		
QY 183	LETTLDVVGKMKFDRGYLSPEVYTNPEKVCLENDPYILCNEKITSMDMLPILBOYAK 242		
DB 182	LENELSVGEMQDRGYISPYFINQONNSAELEHFFILVDKIKATIRDMLSVLEAVAK 241		
QY 243	VNRPILLIADVEGEALATLVNKKLGALQVAVKAPGGERKKAMLEDAITLTGEAIF 302		
DB 242	SGRPILLIADVEGEALATLVNKNMGIVKAPGCDRRKAMLODAITLTNOVIS 301		
QY 303	EDRGICILENVYLSGTAKRVVTDKENTTVVDAGKSEDIKARVKOIRAOIEETSSDYDR 362		
DB 302	EEIGTLETAASLESIGTAKRIYTKENTTIIIDEGAKTEINAKIAOIRAOIEETSSDYDR 361		
QY 363	EKLQERLAKVGVAVIHGAATEMEKKEDVEDALNATRAVEGCIYPCGCTAFVRS 422		
DB 362	EKLQERLAKVGVAVIHGAATEMEKKEDVEDALNATRAVEGCIYPCGCTAFVRS 422		

```

Db      362  EKLDGRVAKLGGVAVINVGATATEIEMKEKARVEDALHATRAAVEEIVAGGVALLRA 421
Qy      423  IKVLDDIKPADDEIAGLNIIRSLSEEPRLQI AANAGYEGSIVYEKREPKDFGCFNAAS 482
Db      422  OKVLITGLGDADNDMGINILKRAIESPLROI VANAAGYESSIVYKVAEHKDNFCFNAT 481
Qy      483  GEYEDLIAGYIDPRKVTIRIALQNNASVASLLTTECIAEKPEPKKMPRG----GGM 538
Db      482  GQYDQMWEMGILDPKTVRTALQNNASVRSIMLTTECYAVDL--FKKDEGAGAGADMGGM 539
Qy      539  GGMGMDGM 547
Db      540  GGMGGMGM 548

```

```

RESULT      2
JN0509
heat shock protein groEL (clone Rhz A) - Rhizobium meliloti
N:Alternate names: chaperonin groEL protein
C:Species: Rhizobium meliloti
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 26-Aug-1999
C:Accession: JN0509
R:kusagawa, E.; Gupta, R.S.
ie 128, 67-75, 1993
file: cloning and characterization of multiple groEL chaperonin-encoding genes in Rhizobium meliloti
A:Reference number: JN0509; MUID:93231539
A:Accession: JN0509
A:Molecule type: DNA
A:Residues: 1-545 <RUS>
A:Cross-references: GB:994192; NID:g152223; PIDs:AAA26285.1; PID:g1522235
C:Comment: This protein plays a role in protein folding and in the extracellular transp
C:Genetics:
A:Gene: groEL
C:Superfamily: chaperonin groEL
C:Keywords: heat shock; molecular chaperone; stress-induced protein

```

[illegible]

QY	479	NAASEYEDLTKACVYIDPKKVTRTALONASVASLITTECAINAKRPEPKMDPMGCGM	538
DB	479 <td>NAQCEIYEDMTLAKMGIIDPVKVVYRTALODASVASLITTEMTAEI--PKKDAAPMPGGM <td>536</td> </td>	NAQCEIYEDMTLAKMGIIDPVKVVYRTALODASVASLITTEMTAEI--PKKDAAPMPGGM <td>536</td>	536
QY	539	CGMGGMGDM	547
DB	537	GGMGGMGDM	545

RESULT 3
R82783 60kDa chaperonin Xf0615 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_rev18on 20-Aug-2000 #text_change 02-Sep-2000
C:Accession: F82783
R:Anonymous. The Xylella fastidiosa Consortium of the Organization for Nucleotide S
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: F82783
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-547 <STM>
A:Cross-references: GB:R82003907; GB:R82003849; NID:99105484; PIDN:AAF83425.1; GSPDB:
A:Experimental source: strain 9a5c
R:Stimpson, A.J.G.; Reznach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga,
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carr
as-Neto, E.; Docena, C.; El-Dorty, H.; Facchinati, A.P.; Ferreira, A.U.S.
Submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.;
J.D.; Junqueira, M.L.; Kemper, E.L.; Klatsjma, J.P.; Krieger, J.E.; Kuramae, E.E.;
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Marti
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeira
Rodrigues, V.; Rosa, A.J.; de M., de Rosa Jr., V.E.; de Sa, R.G.; Sanceli, R.V.; Sa
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore,
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0615
C:Superfamily: chaperonin groEL

[illegible]

QY	361	DREKIOEPLALVGVNAVYIHGAATETEMKEKKDQVEBALNATRAAVEEGIVPGGTAFAV	420
Db	361	DREKIOEPLALVGVNAVYIHGAATETEMKEKKDQVEBALNATRAAVEEGIVPGGTAFAV	420
QY	421	RSIVLIDJIKAPDDELGLNLTIRSLPEPLROIANAANGYEGSIVFEFVREPKDGFNA	480
Db	421	RSIVLIDJIKAPDDELGLNLTIRSLPEPLROIANAANGYEGSIVFEFVREPKDGFNA	480
QY	481	ASGEYEDLIKACVTDPPKVTPLALONAAVSALLITTECAIAEKREPKDKMPM--GGG	537
Db	481	ATGCEGDMVNIIGLDPKVTKRSALONAAISIGLMTTEAMVAE--APKKDEPTPAACGG	538
QY	538	MCGMGMD	545
Db	538	MCGMGMD	546

ULT 4
 ,073
 chaparonin groEL - Chromatium vinosum
 C:Species: Chromatium vinosum
 C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 26-Aug-1999
 C:Accession: B47073
 R:Ferreira, R.G.; Soncini, F.C.; Viale, A.M.
 J. Bacteriol. 175, 1514-1523, 1993
 A:Title: Cloning, characterization, and functional expression in *Escherichia coli* of che
 A:Reference number: A47073; MUID:93186721
 A:Accession: B47073
 A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-546 <FER>
 A:Cross-references: GB:899443; NID:g145007; PIDN:AAA2319.1; PID:g145009
 A:Note: sequence extracted from NCBI backbone (NCBI:n126968, NCBI:p126970)
 C:Superfamily: chaparonin groEL
 C:Keywords: molecular chaparon

Query Match	68.4%	Score 1852.5;	DB 2;	Length 546;
Best Local Similarity	68.2%	Pred. No. 7.3e-83;		
Matches 374; Conservative	70;	Mismatches 101;	Indels 3;	Gaps 2

[illegible]

OY 481 ASGEEDLIRKGVLDPPKKVTFIALONANASVALITTECAIAEKPCKMDP -MPGGMG 539
| | | | | : | | | | | : | | | | | : | | | | |
Db 481 ANGEYGDMEVEGILDPRTKTRSAIÖNSCSVAGILMITEAMTAD--ETKDDPAFAPMPGGMG 538

OY 540 GMSGMDGM 547
| | | | |
Db 539 DMGGMGMM 546

RESULT 5
A41468
60K heat shock protein htpB - Legionella pneumophila

C:\Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999

R; Hoffman, P.S.; Houston, L.; Butler, C.A.

A:Title: *Legionella pneumophila* htpAB heat shock operon: nucleotide sequence and expression
A:Reference number: A1468; MUID:90382960
Intect. Immun. 36, 3360-3367, 1990

A;Accession: A41408
A;molecule type: DNA

```
A;Residuals: 1-550 <HOF>
A;Cross-references: GB:M31918
```

R; Sampson, J.S.; O'Connor, S.P.; Holloway, B.P.; Plakaytis, B.B.; Carlone, G.M.; Ma Infect. Immun. 58, 3154-3157, 1990

A;Title: NUCLEOTIDE SEQUENCE OF H
A;Reference number: A41472; MUID:5

A;Accession: A41472
A;Molecule type: DNA

A;Residues: 3 188, 'S', 190-195, 'LIAVH', 202-34/, 'T', 349-350 <SAMS>
A;Cross-references: GB:M91673; GB:M35149; NID:q149689; PIDN:AAA25298.1; PID:q149690

C;Genetics: <http://www.genetics.org>
A;Gene: <http://www.ncbi.nlm.nih.gov/Genbank/>

C; Superfamily: chaperonin groEL

Query Match	68.1%	Score 1845	DB 2	Length 550
Best Local Similarly	67.3%	Pred. No. 1.7e-82		
Matches 367; Conservative	80;	Mismatches 96;	Indels 2;	Gaps 1

QY 3 SKEILFDAKAREKLSRGVDKLANAVKVTLCGPKGRNVIEKSFSPVITTKDGVSAKEIEL 62

Db 4 AKELRFCDARLQMLAGVNALADAVQVTMGPRGRNVLEKSYGAPVTTKDGSVAKEIEF 63

0Y 63 EDKFNMGAMVKEVAPKTSDIAGDGTTTATVLAQAIYREGVKLVAAGRNPMAIKRGIDK 122

Db 64 EHRFMNMGAMVKEVASKTSDTAGDGTATTATVLARSILVEGHKAVAAGMNPMDLKRIDK 122

QY 123 AVVAVTKEISDITKPTRDQKEIAQVGTISANDTTIGNIIAEAMAKVGKGVITVEEAKG 18

Db 124 AVLAVTKLQAMSKPCKDSKALIAQVGTISANSDEAIGALIAEAMEKVGKEGVITVEDGNG 18

QY 183 LETTLDVVEGMMFDRGYLSPLYVTNPPEKMVCELDNPYILCNEKKITSMKMDLPILEQVAK 24

Db 184 LENELYVEGMDRGYISPYFINNQNMSCLEHPILLVDKKVSSIREMLSVLEGVAK 24

243 VNRPLLI AEDVEGEALATLVV NKL RGAIQVVA VKAPGFGERRKAMLEDAI LTGGEAIF 300

Db 244 SGRPLLIADVEGEALATLVVNNMRGIKVCVAKAFGFGDRKKAMLDIAILTKGVIS 30

303 EDGKLENSLSSLTAKRVVIDKENTTIVDGAGKSEDIKARVKQIRAOIETSSDYDR 36

Db 304 EEIGKSLGATLEDLSAKRIVTKENTTTIDGEGKATEINARIAQIRAQMEETTSYDR 36

QY 363 EKIQLERLAKLVGVAVIHVGATETEMKEKKDRVEDALNATRAAVEECIVPGGTAFVRS 422

Db 364 EKLQERVAKLAGVAIVKGAATEVEMKEKKARVEDALHATRAAVEEGIVAGGVALIRA 422

QY 423 IKVLDDIKPADDELAGLNIIRSLLEPLRQIAANAGYEGSIVEKVKREPKDGFNAAS 48

Db 424 QKALDSLKGDDDDQNGINILRRRAIESPMRQIVTNAGYEASVVVNVKVAEHKIDNYGFNAAT 48

QY 483 GEYEDLIKAGVIDPKKVTTRIALQNASVASLLLTTECAIAEKPEPKKMDMPGGGGMG 54

Oy 1 MASKEILFDKARKRKISRCVDKILANAVKYLGPGRNVVIEKSFSPVITRKDGVSAKEI 60
 Db 1 MAAEVEFGDSARRKKLVGVNVLADAAVKATLGGKGRNVVLDDKSFQPTITKDGVSAAKEI 60
 Oy 61 FLEBKFPMNMAQMWKEVAPKPTSDIAGCGTTATVLAQATIRECVKILVAGGRNMAIKRGI 120
 Db 61 FLKMKPFNMMAQLVKDVASAKANAAGDGTITATVLAQATVIBELKVAAGMNMMDIKRGI 120
 Oy 121 DKAVAVATKELSDITKPTPRDQKEIAOVGTISANSPTTIGNITAEAMAKYKGGVITVEEA 180
 Db 121 DKATVAIVAAOLKEIAKPCADTKAIAQVGTISANSDESIOQIIAEAMEXKKEGVITVEEC 180
 Oy 181 KGIETITLVVVEGKMPKRGYISPYFVYNPREKMCGLDNPYILCNKKITSMDMLPILEOV 240
 Db 181 SGLNELSVVEGQFDRGYLSPYFVVKPDTMAELDSPLLLVDDKISINIREMLPYLEAV 240
 Oy 241 AKNRPRLIIAEVVEGEALATLVNKLQALOVAAKAPGFERRRAMLEDIAITLGGGA 300
 Db 241 AKAGRPRLIIAEVVEGEALATLVNKNMRGIYVAAVKAAPGFERRRAMLEDIAITLGGGV 300
 Oy 301 IFEDRKIKLEBNVSLSLGTAKRNVIDKENTTIVDGAKSSEDIKARYKQIRAOIEETSSDY 360
 Db 301 ISEEVVLSTFCGATLLEHLGNMAKHVIVNKENTTIIIDGAGVADIEARVLQIRKQIEETSDY 360
 Oy 361 DRKILQERLAKLVGCVAAVHGAATETMEKKKDVVEDALNMTRAVEEGVYPGCGTAPV 420
 Db 361 DRKQLQERLAKLVGCVAAVHGAATETMEKKKAREREDALNMTRAVEEGVYPGCGVAVL 420
 Oy 421 RSIVKVLDDIKPADDELAGLNTIRKSLERPLQIQAANAGYEGSVIYEVREPKDGFGRNA 480
 Db 421 RALQALIEGLKQNEHQNVCIALLRRAVESPLQIQAANAGDESVYVDDKQCSNGYGRNA 480
 Oy 481 ASGEVEDLKKAVIDPKVTRIALDNMAVSALLITTECAIAEKPEPKDKMPGCG--GM 538
 Db 481 ATGVIGDMTEMGILDPKAVTRISALQAASIGCLMITTEMAVAEIVE--DKPAMGMPDM 537
 Oy 539 GGMGMDGM 547
 Db 538 GGMGMDGM 546
 RESULT 11
 S22347
 groEL - Brucella abortus
 C.Species: Brucella abortus
 C.Date: 22-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 26-Aug-1999
 C.Accession: S22347
 R.Gor, D., Mayfield, J.E.
 Biochim. Biophys. Acta 1130, 120-122, 1992
 A.Title: Cloning and nucleotide sequence of the Brucella abortus groE operon
 I.Sequence number: S22346; MUID:92182006
 .cession: S22347
 n.status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-546 <GOR>
 A.Cross-references: EMBL:M82975; NID:q144109; PIDN:AAA22997.1; PID:q144111
 C:Superfamily: chaperonin groEL

Query Match	67.3%	Score 1923;	DB 2;	Length 546;
Best Local Similarity	68.9%	Pred. NO. 2e-81;		
Matches 378;	Conservative 64;	Mismatches 99;	Indels 8;	Gaps 5
OY	1	MAKSEILFDPAKRRKRLSRCVDKLANNAVKVTLCPKGNVNIIEKSFSGSPVITKOGVSVAKEI	60	
DB	1	MAAKVYKCRQIRAREKMLKQVDLLAAVAVYTLDPKGNVNIIEKSFSGSPVITKOGVSVAKEV	60	
OY	61	ELEDKFEENKGAOMVEVAPKTSIDAGDGTATVLAQAIREGVKVLAAGRPMMAIKREI	120	
DB	61	ELEDKFEENKGAOMLEEVASKTNDPTAGDGTATVLAQAIYOGKARAVAAAGNMPMDLKRRI	120	
OY	121	DKAAVAVAKLEISDITKPTPRDOKEIQAQVCTISANSDDTTIGNIIIAEKMAKVGKGVLTVEEA	180	

```

Db      121 DLAVNEVMAELLKKAKKINTSEEVAOVGTISANNAEATGKMAEAMOKVNECVITVEEA 180
QY      181 KGLETTLDVVEGKNEFDRGYLSPYFTYNPEKMCYCLDNPHYILCNEKRTKTSKKMDPLITEOY 240
Db      181 K7AETELEVEVEGQFDRGYLSPYFTYNPEKVADEDEAYILHHEKKSJNIOALPLVLEAA 240
QY      241 AKVNPRLIIAEDVEGALATLVNKLGCALQVAAVNAKPGFERKAMLEDIAILNGEX 300
Db      241 VQTSKPLIIAEDVEGALATLVNKLKRGJLKAIAVAPGQDRKMLEDIAILNGOY 300
QY      301 IEFDRGRIKENYSLSSLGTAKEVYIDKENTTIVDGAGSEDIKARVQIARQIETSSDY 360
Db      301 ISEDLGKLESVTLMLGRAKVSSISKENTTIVDGAGKAEIDARVOQIQOIEETSSDY 360
QY      361 DREKLOERLAKIVGVAVYIHGAATETEMKREKRVEDALNATRAAVEGIVGGTAVY 420
Db      361 DREKLOERLAKIAGVAAVIRGVGAIEVEVKRORVDALNATRAAVEGIVAGSTALL 420
QY      421 R-SIKVLDDIKPADDELAGINIIIRSLSEELRQIANAGYEGSIVAEVYKRE-PKRGPF 478
Db      421 RASTKI--TAKVYNAVDQENGINIVRRAIQAPARITTNAAEEASVYIGKILETSETFCY 478
QY      479 NAASEYEDLAKAGVIDPKKATYRIALONAAASVASILITTECAIAEKPPEPKOMPNG--G 536
Db      479 NNANGEYGDLSLGIQDPVKVYVTRALQNAASVAGLITTEAMIAEL--PKKDAAPACMPG 536
QY      537 GNGGNGMD 545
Db      537 GNGGNGMD 545

```

```

RESULT      12
S35309
heat shock protein groEL2 - Bradyrhizobium japonicum
C:Species: Bradyrhizobium japonicum
C>Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 26-Aug-1999
C:Accession: S35309
R.Fischer, H.M.; Babst, M.; Kaspar, T.; Acuna, G.; Arizoni, F.; Hennecke, H.
EMBO J. 12, 2901-2912, 1993
A:Title: One member of the groEL-like chaperonin multigene family in Bradyrhizobium
A:Reference number: S35308; MUID:93327780
A:Molecule type: DNA
A:Residues: 1-550 <FIS>
A:Cross-references: EMBL:Z22604; NID:g312977; PIDD:CAA80318.1; PID:g312979
C:Genetics:
A:Gene: groEL2
C:Superfamily: chaperonin groEL
C:Keywords: molecular chaperone

Query Match          67.3%   Score 1823; DB 2; Length 550;
Best Local Similarity 66.7%; Pred. No. 2e-81;
Matches 368; Conservative 81; Mismatches 93; Indels 10; Gaps 4;

OY    1 MASKEILFDARAKRELTSQVDKLANAVVTLTGPGGRNVIRIKSGSPVITODGSVAKEI 60
       |::|::| |::|::| |::|::| |::|::| |::|::| |::|::| |::|::| |::|::|
DB    1 MSAAKVRKGVDARDMLRGVDILHNAVKVTLGPGGRNVLDKSGCAPRITDGVTAKEI 60
       |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

OY    61 ELEDFKFENMGAMQWKEVAPKTSDIAGDGTATVLAQAIVREGVKLVACGRNPAIRGI 120
       |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    61 ELEDFKFENMGAMQWREVASKSADAAAGDGTATTVALAIIVREGAKSVAAAGNPMDLRGI 120
       |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

OY    121 DKAVAAYVKELSDITKPTRDKETIAOVGTITSANSDTTIGNIIAEAMAKVGKGIIVEEA 180
       |::|::| |::|::| |::|::| |::|::| |::|::| |::|::| |::|::| |::|::|
DB    121 DMAEAAVADVADLVKKSKKWTSSNEEIIAOVGTISANGDAIEIGKFTISPMKKVGNEGVIIVEEA 180
       |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

OY    181 KGLETTILDVVEGMKFRDGYSLEFVTPNPEKKWVCLENDPIYCNCKKITSNKMDLPILLEOV 240
       |::|::| |::|::| |::|::| |::|::| |::|::| |::|::| |::|::| |::|::|
DB    181 KSLEIELEVVEGMQFPDRGYISPLYFTVNADKRVRKMDDAIYVLINKKLSQLNELLPILLEAV 240
       |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

OY    241 AKVNRPLLTIIAEDVGEGLATLVVNLKGALOVVAVKAPGFGERRKMALEDIALLTGGEA 300

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2001, 17:18:44 ; Search time 57.93 Seconds

(without alignments)
324,046 Million cell updates/sec

Title: US-09-077-574a-2

Perfect score: 2708

Sequence: 1 MASKEILFPAKAREKLSRGV.....KDMMPGCGMGCMGMDMY 548

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

arched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SWISSProt_39:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1872.5	69.1	546	1 CH60_BURCE	Q92FE0 burkholderi
2	1868.5	69.0	546	1 CH60_BURVI	Q92FD8 burkholderi
3	1867.5	69.0	547	1 CH60_BARKH	O33963 bartonella
4	1867	68.9	545	1 CH61_RHIME	P35469 rhizobium m
5	1866	68.9	551	1 CH60_AMOPS	P26004 amoeba prot
6	1864.5	68.9	547	1 CH60_BAROU	O33964 bartonella
7	1852.5	68.4	546	1 CH60_CHRYI	P31293 chromatium
8	1845	68.1	545	1 CH60_RHCCA	P95678 rhodobacter
9	1844.5	68.1	546	1 CH60_RHILV	P34939 rhizobium l
10	1843.5	68.1	547	1 CH60_BORPE	P48210 bordetella
11	1843.5	68.1	552	1 CH60_COXBU	P15421 coxiella bu
12	1834.5	67.7	546	1 CH61_RHOSH	P20110 rhodobacter
13	1834.5	67.7	552	1 CH60_PSEST	O33500 pseudomonas
14	1833.5	67.7	539	1 CH61_BRAJA	P77829 bradyrhizob
15	1828.5	67.5	544	1 CH60_AGRUV	P30779 agrobacteri
16	1825.5	67.4	547	1 CH60_PSEAE	P30718 pseudomonas
17	1823	67.3	546	1 CH60_BRUAB	P25967 brucella ab
18	1818	67.1	549	1 CH62_BRAJA	P35861 bradyrhizob
19	1816.5	67.1	541	1 CH63_RHIME	P35470 rhizobium m
20	1816	67.1	542	1 CH62_RHIME	P35470 rhizobium m
21	1814.5	67.0	544	1 CH60_NEIMA	P43005 neisseria m
22	1808.5	66.8	544	1 CH60_NEIMA	P43005 neisseria m
23	1805.5	66.7	543	1 CH60_BARRA	P35635 bartonella
24	1802.5	66.6	544	1 CH60_NEIGO	P28842 neisseria g
25	1800.5	66.5	547	1 CH60_LEGPN	P26878 legionella
26	1799	66.4	546	1 CH60_LEGMI	P26878 legionella
27	1794.5	66.3	544	1 CH60_NEIFL	P48215 neisseria f
28	1789	66.1	545	1 CH63_BRAJA	P35863 bartonella
29	1786	66.0	545	1 CH60_AQUAE	O67943 aquifex aco
30	1780.5	65.7	544	1 CH60_BACHD	O50303 bacillus ba
31	1774.5	65.5	545	1 CH60_PSEPU	P48216 pseudomonas
32	1769.5	65.3	544	1 CH60_PSEPU	P48216 pseudomonas
33	1765	65.2	546	1 CH60_LEPIN	P35468 leptospira

34	1760.5	65.0	539	1 CH60_BACST	Q07201 bacillus st
35	1759	65.0	548	1 CH60_HAEIN	P43733 haemophilus
36	1758.5	64.9	545	1 CH60_CAUCR	P48211 caulobacter
37	1755	64.8	545	1 CH60_CAMEE	O69289 calyobacter
38	1750.5	64.6	546	1 CH60_ACTPL	P94166 actinobacill
39	1750	64.6	550	1 CH60_YERN	P48219 yersinia en
40	1747	64.5	547	1 CH60_SALVI	P48217 salmonella
41	1745	64.4	547	1 CH60_PASMU	O59687 pasteurilla
42	1743.5	64.4	543	1 CH60_BACSU	P28598 bacillus su
43	1742.5	64.3	545	1 CH60_PAREE	O92462 paracoccus
44	1740.5	64.3	546	1 CH60_ACTPC	P46398 actinobacill
45	1739	64.2	547	1 CH60_HAEUD	P31294 haemophilus

ALIGNMENTS

RESULT 1	CH60_BURCE	STANDARD:	PRT: 546 AA.
ID	CH60_BURCE		
AC	O92FE0:		
DT	30-MAY-2000 (Rel. 39, Created)		
DT	30-MAY-2000 (Rel. 39, Last sequence update)		
DT	30-MAY-2000 (Rel. 39, Last annotation update)		
DE	60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL, PROTEIN).		
GN	MOPR OR GROEL.		
OS	Burkholderia cepacia (Pseudomonas cepacia).		
OC	Bacteria; Proteobacteria; beta subdivision; Burkholderia group;		
OX	NCB1_TaxID=292;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=NCIC 10744;		
RA	Zysk G., Spletstoeser W.D., Neubauer H.;		
RT	"Nucleotide sequence comparison of the groE operon of Burkholderia		
RT	spp.";		
RL	Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.		
CC	- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND		
CC	PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS		
CC	CONDITIONS (BY SIMILARITY).		
CC	- 1 SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF		
CC	7 SUBUNITS (BY SIMILARITY).		
CC	- 1 SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation-		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@sib-sib.ch).		
CC	-----		
CC	EMBL: AF104907; AAC79087.1; -		
DR	InterPro: IPR001844; -		
DR	InterPro: IPR002423; -		
DR	Pfam: PF00118; CPN60_TCP1; 1.		
DR	PRINTS: PR00298; CHAPERONIN60.		
DR	PRINTS: PR00304; TCOMPLEXCP1.		
DR	PROSITE: PS00296; CHAPERONIN5; CPN60; 1.		
KW	Chaperone; ATP-binding.		
SEQ	SEQUENCE 546 AA; 56980 MW; 174B9934345E7315 CRC64;		

Query Match 69.1%; Score 1872.5; DB 1: Length 546;
Best Local Similarity 68.4%; Pred. No. 1e-79;
Matches 374; Conservative 75; Mismatches 95; Indels 3; Gaps 2;

QY	1 MASKEILFPAKAREKLSRGVDTLANAVKTLCPKGNVYKSPGSPVTTKGVSAKEI 60	:
DB	1 MAKADVFQDSRSKRWESVNLANVKTTLDPKGNVLENSFGCPYTKGVSAKEI 60	:
QY	61 ELEDKFNKGAQVKEVAPKTSIDAGDGTATVLAQATYRGVKKLVAAGRPMIAKRI 120	:

```

Db 61 ELKDKLQNMGAOMVEVASKTSDNAAGDGTATVLAQSIIVEGKMKYVASGNMPLDKRGI 120
Oy 121 DKAAVAATKELSDITKPRDOKETIAOQVCTISANSDDTTGNIIEAMAVGCVITVEEA 180
Db 121 DKAAVAATKELSDITKPRDOKETIAOQVCTISANSDDTTGNIIEAMAVGCVITVEEA 180
Oy 181 KGLETTLDVVEGKMFDRGYLSPEYFVTPNEKMYCELDNPYILCNEKKITSMKMDLPITLEV 240
Db 181 KGLETTLDVVEGKMFDRGYLSPEYFVTPNEKMYCELDNPYILCNEKKITSMKMDLPITLEV 240
Oy 241 AKVNRPLLIIEADVEGEALATLVNNKLGALQVAVAKAPGCGERRKAMLEDIATITGGEA 300
Db 241 AKVNRPLLIIEADVEGEALATLVNNKLGALQVAVAKAPGCGERRKAMLEDIATITGGEA 300
Oy 301 IFEDRGITLENVSLSSICTAKRVVIDKENTTIYDCAKSEDIKARVQIRAOIEETSSDY 360
Db 301 IFEDRGITLENVSLSSICTAKRVVIDKENTTIYDCAKSEDIKARVQIRAOIEETSSDY 360
Oy 361 DREKLOERLAKLVGVAIVHVGATETEMKCKDRVEDALNATRAAEEGIVPGGTAFFV 420
Db 361 DREKLOERLAKLVGVAIVHVGATETEMKCKDRVEDALNATRAAEEGIVPGGTAFFV 420
Oy 421 RSIVLDDIKPADDELACLNITRRSLEELPLQJIAANAGYEGSIIVEKREPKDGFENA 480
Db 421 RSIVLDDIKPADDELACLNITRRSLEELPLQJIAANAGYEGSIIVEKREPKDGFENA 480
Oy 481 ASGEYEDLIKAGVIDPKKVTIRIALQNAASVSLTTTECAIAEKPERKKDMPMGCGMG 540
Db 481 ASGEYEDLIKAGVIDPKKVTIRIALQNAASVSLTTTECAIAEKPERKKDMPMGCGMG 540
Oy 541 MCGMDGM 547
Db 539 MCGM-GM 544

RESULT 2
CH60_BURVI STANDARD: PRT: 546 AA.
ID CH60_BURVI
AC 092FDB:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN).
GN MOPA OR GROEL.
OS Burkholderia vietnamiensis.
OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group.
OC Burkholderia.
OX NCBI_TaxID=60552;
RN [1]
RP SEQUENCE FROM N.A.
RT STRAIN=DSM 11319;
"Nuclotide sequence comparison of the groE operon of Burkholderia spp."
SUBMITTED (NOV-1998) TO THE EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CC CONDITIONS (BY SIMILARITY).
CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
CC 7 SUBUNITS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF104908; AAC79089.1;
CC DR InterPro: IPR001844;
CC DR InterPro: IPR002423;

```

```

DR Pfam: PF00118; cpn60_TCP1.1.
DR PRINTS: PR00298; CHAPERONIN60.
DR PRINTS: PR00304; TCOMPLEXTCP1.
DR PROSITE: PS00296; CHAPERONIN_CP60.1.
KW Chaperone; ATP-binding.
SQ SEQUENCE 546 AA; 57010 MW; 47FA893959221218 CRC64;

Query Match 69.0%; Score 1868.5; DB 1; Length 546;
Best Local Similarity 68.2%; Pred. No. 1.6e-79;
Matches 373; Conservative 75; Mismatches 96; Indels 3; Gaps 2;

Oy 1 MASKEITLFDKAREKLSGVCKLNAVYKVTGPGGRNVITKSGSPVITDGVSKKEI 60
Db 1 MAKDQVFGDSARSKMVEGNVILNAAYKVTGPGGRNVITKSGSPVITDGVSKKEI 60
Oy 61 ELKDKFENMGAMQKVEVAPKTSIDAGDGTATVLAQSIIVEGKMKYVASGNMPLDKRGI 120
Db 61 ELKDKLQNMGAOMVEVASKTSDNAAGDGTATVLAQSIIVEGKMKYVASGNMPLDKRGI 120
Oy 121 DKAAVAATKELSDITKPRDOKETIAOQVCTISANSDDTTGNIIEAMAVGCVITVEEA 180
Db 121 DKAAVAATKELSDITKPRDOKETIAOQVCTISANSDDTTGNIIEAMAVGCVITVEEA 180
Oy 181 KGLETTLDVVEGKMFDRGYLSPEYFVTPNEKMYCELDNPYILCNEKKITSMKMDLPITLEV 240
Db 181 KGLETTLDVVEGKMFDRGYLSPEYFVTPNEKMYCELDNPYILCNEKKITSMKMDLPITLEV 240
Oy 241 AKVNRPLLIIEADVEGEALATLVNNKLGALQVAVAKAPGCGERRKAMLEDIATITGGEA 300
Db 241 AKVNRPLLIIEADVEGEALATLVNNKLGALQVAVAKAPGCGERRKAMLEDIATITGGEA 300
Oy 301 IFEDRGITLENVSLSSICTAKRVVIDKENTTIYDCAKSEDIKARVQIRAOIEETSSDY 360
Db 301 IFEDRGITLENVSLSSICTAKRVVIDKENTTIYDCAKSEDIKARVQIRAOIEETSSDY 360
Oy 361 DREKLOERLAKLVGVAIVHVGATETEMKCKDRVEDALNATRAAEEGIVPGGTAFFV 420
Db 361 DREKLOERLAKLVGVAIVHVGATETEMKCKDRVEDALNATRAAEEGIVPGGTAFFV 420
Oy 421 RSIVLDDIKPADDELACLNITRRSLEELPLQJIAANAGYEGSIIVEKREPKDGFENA 480
Db 421 RSIVLDDIKPADDELACLNITRRSLEELPLQJIAANAGYEGSIIVEKREPKDGFENA 480
Oy 481 ASGEYEDLIKAGVIDPKKVTIRIALQNAASVSLTTTECAIAEKPERKKDMPMGCGMG 540
Db 481 ASGEYEDLIKAGVIDPKKVTIRIALQNAASVSLTTTECAIAEKPERKKDMPMGCGMG 540
Oy 541 MCGMDGM 547
Db 539 MCGM-GM 544

RESULT 3
CH60_BURVI STANDARD: PRT: 547 AA.
ID CH60_BURVI
AC 033963; 087267;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN) (HEAT SHOCK PROTEIN
DE 60).
GN MOPA OR GROEL.
OS Bartonella henselae (Rochalimaea henselae).
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales group;
OC Bartonellaceae; Bartonella.
OX NCBI_TaxID=38323;
RN [1]
RP SEQUENCE FROM N.A.
RT STRAIN=ATCC 49882;
RA MDLHE-97419519; PubMed-9274034;
RA Haake D.A., Summers T.A., McCoy A.M., Schwartzman M.;
RT "Heat shock response and groEL sequence of Bartonella henselae and

```

```

RT Batonella quintana."
RL Microbiology 143:2807-2815(1997).
RN [2]
RP SEQUENCE OF 1-543 FROM N.A.
RC STRAIN-HOUSTON-1;
RA Marston E.L., Sumner J.W., Regnery R.L.;
RT "Evaluation of intraspecies genetic variation within the 60 kDa heat
RT shock protein (groEL) gene of Bartonella species: a new phylogenetic
RT analysis tool."
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-408 FROM N.A.
RC STRAIN-HOUSTON-1;
RA Sumner J.W., Nicholson W.L., Massung R.F.;
RT "PCR amplification and comparison of nucleotide sequences from the
RT groEL heat shock operon of Escherichia species."
J. Clin. Microbiol. 35:2087-2092(1997)
CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CC CONDITIONS (BY SIMILARITY).
CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
CC 7 SUBUNITS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
CC .....
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on ways
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb.sib.ch).
CC
DR EMBL: U78514; AAB69094.1;
DR EMBL: AF014829; AAB04238.1; ALT_INIT.
DR EMBL: U96734; AAB65637.1;
DR InterPro: IPR001844;
DR InterPro: IPR002423;
DR Pfam: PF00118; cpn60_TCP1; 1.
DR PRINTS: PRO0298; TCPEPROMIN60.
DR PRINTS: PRO0304; TCOMPLEXTCP1.
DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
KW Chaperone; ATP-binding; Heat shock.
SQ SEQUENCE 547 AA: 57625 MW: EDE2D5D66D5B10D2 CRC64:

Query Match 69.0% Score 1867.5; DB 1; Length 547;
Best Local Similarity 70.1% Pred. No. 1.7e-79;
Matches 385; Conservative 61; Mismatches 96; Indels 7; Gaps 4;

QY 1 MASKEILFDAAREKLSGVDKLANAVVLTGPGKGRNVIEKSPFSPIYTKDGVSAKEI 60
  ||||| : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 MAKEEVKGRREARRELLRGVDILANAVVLTGPGKGRNVIDKSPFAPRITKGVSAKEI 60

QY 61 ELEDKFKMGQAKYEVAKPTSDIAGDGTATYVLAQAIYREGVKVLAAGRNPAAIKRGI 120
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 ELEDKFKEMGAQMLREVASKTNDIAGDGTATYVLAQAIYREGVKVLAAGMPMDIAKGI 120

QY 121 DKAVAAVKKELSDITKPTPRDQKEIAQVGTISANSPTTIGNIIAEMAAVVGKGVITVEEA 180
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 DAAVDEVANALFKKAKKIQTSAEIAQVGTISANGAAEIGKMTADMEVVGNEGVIITVEEA 180

QY 181 KGLLETTLDVVEGCMKFDRCYLSPEYVNTPEKMKVCELDNPFILCNEKKITSMKMDLPILDEV 240
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 KTAETEELEVVGCMQDRDRLSPYFVTNAKKNVADLDQDYILHKKKLSNLGSLPVLNAV 240

QY 241 AKVNRPLIIIDEDVEGELATLVNKKLGALQVAVKAPGFEGERKKAMLEDIAITLTGGEA 300
  : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :
DB 241 VQSGKPLIIIDEDVEGELATLVNKKLGGLKIAAVKAPGFCDDRKKAMLEDIAITLTGCV 300

QY 301 IFFDRCIKLENVSSLSCTAKRNVYIDKENTTIVDAGKSESDIKARVKQIRAOIEETSSDY 360
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 301 ISEDVGIKLENVYTDMLDIAKRAKVNISKENTTIIDAGOKSEINARVNOIKVQIEETTSY 360

```

```

OY 361 DREKIOEFLATLYGVANVYIHGAAETEMKEKKORVEDALNATRAAVEGIVPGGTAFFV 420
Db 361 DREKIOEFLATLYGVANVYIHGAAETEMKEKKORVEDALNATRAAVEGIVPGGTAFFV 420
OY 421 RSIKVLDDIKPADDELGLNIRISLEEPILJOIANANGVEGSIIVEYVERPK--DGFEN 479
Db 421 RAANAL-IVKGSNPQOEGINIVRRALDAPARQITNNGEFAIIVGKLENNADTFGYN 479
OY 480 AASGEVDIKAGVIDPKKVTRIALONASVASILLTTTECAIAEKPEPKDM--PMGG 536
Db 480 TATGFEGLLALGIIVDPKVVRSALQONASIASILLITTEAMVAE--VPKDTPVPMGG 537
OY 537 GMGGMGMD 545
Db 538 GMGGMGMD 546

RESULT 4
CH61_RHIME STANDARD: PRF: 545 AA.
AC P35469.
DT 01-JUN-1994 (Rel. 29, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 60 KDA CHAPERONIN A (PROTEIN CPN60 A) (GROEL PROTEIN A).
CN GROEL OR GROEL-A.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_Taxid:382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-1021;
RX MEDLINE-95231539, Pubmed-80971779;
RA Rusanawan E., Gupta R.S.;
RT "Cloning and characterization of multiple groEL chaperonin-encoding
RT genes in Rhizobium meliloti.";
RL Gene 126:67-75(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-1021;
RA Ogawa J.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-1021;
RA Galibert F., Capela D., Hubler-Barloy F., Gaius M., Batut J.,
RA Boistard P., Gouzy J., Kahn D., Thebaud P., Goffeau A.,
RA Purnelle B., Pohl T., Bothé G., Schneider S., Portetle D.,
RA Vandenbol M., Puehler A., Becker A., Weidner S.;
RL Submitted (MAR-2000) to the SWISS-PROT data bank.
CC -!- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CC CONDITIONS (BY SIMILARITY).
CC -!- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
CC 7 SUBUNITS (BY SIMILARITY).
CC -!- INDUCTION: BY HEAT SHOCK.
CC -!- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M94192; AAA26285.1; -
DR EMBL; U19726; AAA61955.1; -
DR PIR; JN0509; JN0509.
DR HSSP; P06139; JGRU.
DR InterPro; IPR001844; -

```

DR InterPro: IP002423; .
 DR Pfam: PF00118; cpn60_TCP1.1.
 DR PRINTS: PR00298; CHAPERONIN60.
 DR PRINTS: PR00304; TCOMPLEXTCP1.
 DR PROSITE: PS00296; CHAPERONINS_CPN60.1.
 DR Chaperone: ATP-binding; Heat shock; Multigene family.
 FT Chaperone: ATP-binding; Heat shock (IN REF. 1).
 KM CONFLICT 12 12 A -> G (IN REF. 1).
 SQ SEQUENCE 545 AA; 57687 MW; 4CEE358957E1B45A CRC64;

Query Match 68.9%; Score 1867; DB 1; Length 545;
 Best Local Similarity 70.7%; Pred. No. 1,8e-79;
 Matches 388; Conservative 58; Mismatches 97; Indels 6; Gaps 4;

OY 1 MASKEFLDPAKAREKSLRVDKLANAVKVTLCGKGRNVVIEKSFSGPYITTKOVSAAKEI 60
 DB 1 MAKEEVKFSRAREKMLRVDLADVAVKVLGKGRNVVIDSFGAPRTTKDVSAAKEI 60
 OY 61 ELEDKFNNGAOMKVEAPRTSDIAGDTTATVLAQAIYREGVKLVAGRNPMIAIKRGI 120
 DB 61 FLEDEKFNNGAOMKVEAPRTSDIAGDTTATVLAQAIYREGVKLVAGRNPMIAIKRGI 120
 OY 121 DKAVAATVTELSDTITKPTROKRIAOVGTISANSDDTIGNIIAEAMAKVKGCVITVEA 180
 DB 121 DLAVAEVVDLTLAKAKKINTSDVAGVGTISANGKQIGLDIAEAMQVNEGIVTVEA 180
 OY 181 KGEFTTLDVVEGKMPDRCYLSPTVTPNPEKMYCELDNPYILCNEKKTSMKMDLPLEOV 240
 DB 181 KTAETLEVEGKMPDRCYLSPTVTPNPEKMYCELDNPYILCNEKKTSMKMDLPLEAV 240
 OY 241 AKVNRPLLIADVEGEALATLVNKLRGALOVAVKAPGGERKRAMLEDIAILITGGA 300
 DB 241 VOTGRKLIIADVEGEALATLVNKLRGALOVAVKAPGGERKRAMLEDIAILITGGA 300
 OY 301 IEDRCIKLENNVSLSLGTAKRVVIDKENTTIVDGAKSEDIKARVKQIRAOILEETSSDY 360
 DB 301 ISEDLIKLENNVSLSLGTAKRVVIDKENTTIVDGAKSEDIKARVKQIRAOILEETSSDY 360
 OY 361 DEPKIOERLAKLVGCAVAVHGAATETEMKEKKDRVEDALNMTRAVEEGIVPGGTAIV 420
 DB 361 DEPKIOERLAKLVGCAVAVHGAATETEMKEKKDRVEDALNMTRAVEEGIVPGGTAIV 420
 OY 421 R-SIKVLDDIKPADDELALGNIIRRSLEPLRQIQAANAGYEGSIYVEKVRPEPK-DGFG 478
 DB 421 R-SIKVLDDIKPADDELALGNIIRRSLEPLRQIQAANAGYEGSIYVEKVRPEPK-DGFG 478
 OY 479 NAASGEYEDLIKAGVIDPKKVTTRIALONASVASILLITTECAIAEKPEPKDMPGGM 538
 DB 479 NAOTGEYEDLIKAGVIDPKKVTTRIALONASVASILLITTECAIAEKPEPKDMPGGM 538
 OY 539 GGMGMDGM 547
 DB 539 GGMGMDGM 547
 OY 537 GGMGMDGM 545
 DB 537 GGMGMDGM 545

RESULT 5
 CH60_AMOPS STANDARD: PRT: 551 AA.
 AC P26004;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN).
 GN MOPR OR GROEL.
 OS Amoeba proteus symbiotic bacterium.
 OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
 OC Legionellaceae.
 OX NCBI_TaxID-2728;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ahn T.I., Leeu H.K., Kwak J.H., Jeon K.W.;
 RT "Nucleotide sequence and temperature-dependent expression of groEL
 gene isolated from symbiotic bacteria of Amoeba proteus.";

RL Endocyt. Cell Res. 8:33-44(1991).
 CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
 CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
 CC CONDITIONS (BY SIMILARITY).
 CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
 CC 7 SUBUNITS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).
 CC -----
 DR EMBL: M86549; AAC09381.1; .
 DR HSSP: P06139; 1GRL.
 DR InterPro: IP001844; .
 DR InterPro: IP002423; .
 DR Pfam: PF00118; cpn60_TCP1.1.
 DR PRINTS: PR00298; CHAPERONIN60.
 DR PRINTS: PR00304; TCOMPLEXTCP1.
 DR PROSITE: PS00296; CHAPERONINS_CPN60.1.
 DR Chaperone: ATP-binding.
 SQ SEQUENCE 551 AA; 58332 MW; A25895CBF9585DA0 CRC64;

Query Match 68.9%; Score 1866; DB 1; Length 551;
 Best Local Similarity 67.9%; Pred. No. 2e-79;
 Matches 373; Conservative 74; Mismatches 96; Indels 6; Gaps 2;

OY 3 SKETILFPAKAREKSLRVDKLANAVKVTLCGKGRNVVIEKSFSGPYITTKOVSAAKEI 62
 DB 3 SKETILFPAKAREKSLRVDKLANAVKVTLCGKGRNVVIEKSFSGPYITTKOVSAAKEI 62
 OY 2 AKELRFGCDAROOMLAGVALADRVKATVPGSGRNVLERSGAPVTFTKDVSAKEIEF 61
 DB 2 AKELRFGCDAROOMLAGVALADRVKATVPGSGRNVLERSGAPVTFTKDVSAKEIEF 61
 OY 63 EDKFNNGAOMKVEAPRTSDIAGDTTATVLAQAIYREGVKLVAGRNPMIAIKRGI 122
 DB 63 EDKFNNGAOMKVEAPRTSDIAGDTTATVLAQAIYREGVKLVAGRNPMIAIKRGI 122
 OY 123 AVVAATVTELSDTITKPTROKRIAOVGTISANSDDTIGNIIAEAMAKVKGCVITVEA 182
 DB 123 AVVAATVTELSDTITKPTROKRIAOVGTISANSDDTIGNIIAEAMAKVKGCVITVEA 182
 OY 182 LETTLDVVEGKMPDRCYLSPTVTPNPEKMYCELDNPYILCNEKKTSMKMDLPLEOV 242
 DB 182 LETTLDVVEGKMPDRCYLSPTVTPNPEKMYCELDNPYILCNEKKTSMKMDLPLEOV 242
 OY 243 VNRPLLIADVEGEALATLVNKLRGALOVAVKAPGGERKRAMLEDIAILITGGA 302
 DB 243 VNRPLLIADVEGEALATLVNKLRGALOVAVKAPGGERKRAMLEDIAILITGGA 302
 OY 303 EDRCIKLENNVSLSLGTAKRVVIDKENTTIVDGAKSEDIKARVKQIRAOILEETSSDY 362
 DB 303 EDRCIKLENNVSLSLGTAKRVVIDKENTTIVDGAKSEDIKARVKQIRAOILEETSSDY 362
 OY 362 EELGTSLETPASLESIGTAKRIYTKENTTIIIDEGKATEINRIQIRAOIMETSSDY 361
 DB 362 EELGTSLETPASLESIGTAKRIYTKENTTIIIDEGKATEINRIQIRAOIMETSSDY 361
 OY 423 IKVLDDIKPADDELALGNIIRRSLEPLRQIQAANAGYEGSIYVEKVRPEPKDMPGGM 482
 DB 423 IKVLDDIKPADDELALGNIIRRSLEPLRQIQAANAGYEGSIYVEKVRPEPKDMPGGM 482
 OY 482 QKVLGDKDNDQMGNIITLRAITSPRLQIQAANAGYSSVIVNVAEHKDFGNAA 481
 DB 482 QKVLGDKDNDQMGNIITLRAITSPRLQIQAANAGYSSVIVNVAEHKDFGNAA 481
 OY 539 GGMGMDGM 547
 DB 539 GGMGMDGM 547
 OY 540 GGMGMDGM 548
 DB 540 GGMGMDGM 548


```

RESULT 6
CH60_BAROU STANDARD: PRT: 547 AA.
ID CH60_BAROU
AC 033964;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 60 KVA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN) (HEAT SHOCK PROTEIN 60).
GN MOPR OR GROEL.
OS Bartonella quintana (Rochalimaea quintana).
OC Bacteria: Proteobacteria: alpha subdivision: Rhizobiaceae group:
OC Bartonellaceae: Bartonella.
OX NCBI_TaxID=803;
RN 11;
RP SEQUENCE FROM N.A.
STRAIN-ATCC 51694;
MEDLINE=97419519; PubMed=9274034;
Hlake D.A., Summers T.A., McCoy A.M., Schwartzman W.;
"Heat shock response and groEL sequence of Bartonella henselae and
Bartonella quintana";
Microbiology 143:2807-2815(1997).
RT
CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CC CONDITIONS (BY SIMILARITY).
CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
CC 7 SUBUNITS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U78515; AAB69095.1;
DR InterPro: IPR001844;
DR InterPro: IPR002423;
DR Pfam: PF00118; cpn60_TCP1; 1.
DR PRINTS: PRO0298; CHAPERONIN60.
DR PRINTS: PRO0304; TCOMPLEXTCP1.
DR PROSITE: PS00296; CHAPERONIN60; 1.
DR Chaperone: ATP-binding; Heat shock
SEQUENCE 547 AA: 57610 MW: 86200CF50BF2D6F4 CRC64:

```

```

Query Match 68.9%; Score 1864.5; DB 1: Length 547;
Best Local Similarity 70.1%; Pred. No. 2,4e-79;
Matches 385; Conservative 61; Mismatches 96; Indels 7; Gaps 4;

```

```

QY 1 MASKELLFDKAREKLSGYDKLANAVKVTLPGRNVVTEKSGFVITKDGVSVAKEI 60
DB 1 MAAKEVFGREARELLRGVDILANAVKVTLPGRNVVTEKSGFVITKDGVSVAKEI 60
QY 61 ELEKPFENMGAOMKREVAASSTNDAGCGTTATVLAQAIYREGKVLAAAGNPMALRGI 120
DB 61 ELEKPFENMGAOMKREVAASSTNDAGCGTTATVLAQAIYREGKVLAAAGNPMALRGI 120
QY 121 DKAVVAATKELSDITKPTDQKEIAQVGTISANSDDTIGNIAEAMAKVKGQVITVEEA 180
DB 121 DAAVEEVVGNLFKAKKILQVSAELAQVGTISANCAAEIGKMIADAMKVGSGEYITVEEA 180
QY 181 KGLETVLADVVGKMFDRGYISPYFVITNPEKMCVCELDNPIYLCKNEKTTSMKMLPILEOV 240
DB 181 KGLETVLADVVGKMFDRGYISPYFVITNPEKMCVCELDNPIYLCKNEKTTSMKMLPILEOV 240
QY 241 AKVAPRLIIADEVGEFALNTLVNKLRGALQVAVKAPGGEGRKMLIEDIATLTGGEA 300
DB 241 VOSKPLLIIADEVGEFALNTLVNKLRGALQVAVKAPGGEGRKMLIEDIATLTGGEA 300
QY 301 IFEDRGIKLENVSLISGTAKRNVIDKENTTIYDGAKSDEIKARVQIRAOIEETSSDY 360

```

```

DB 301 ISEDVGIKLENTVLDMLGRAKKVNSKENTTIIDGAKKAEINARVQIRAOIEETSSDY 360
QY 361 DREKLOERLAKLVGCVAVIVGAATETEMKEKKDRVDALNATRAAEEDIVGCGTAFV 420
DB 361 DREKLOERLAKLVGCVAVIVGAATETEMKEKKDRVDALNATRAAEEDIVGCGTAFV 420
QY 421 RSIKVLDDIKPADDELAGNIIRSLSEELPLROIAAAGYEGSIVYKREPK-DGSGFN 479
DB 421 RAANAL-AIKGSNPDOAGINIVRAALQARQATATAGEEAATVYKLENNADTFGCYN 479
QY 480 AASGEEDLIKAGVIDPKVTRIALQNAVASLTLTECAIAEKPEPKDM--PMPG 536
DB 480 TATGFGDLIALGIVDEPKVYSALQNAVASLTLTEAMVAE--VPKKDTMPMPG 537
QY 537 GKGGMGMD 545
DB 538 GKGGMGMD 546

```

```

RESULT 7
CH60_CHRYI STANDARD: PRT: 546 AA.
ID CH60_CHRYI
AC P31293;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 60 KVA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN).
GN MOPR OR GROEL.
OS Chromatium vinosum.
OC Bacteria: Proteobacteria: gamma subdivision: Chromatiaceae:
OC Allochromatium.
OX NCBI_TaxID=1049;
RN 11;
RP SEQUENCE FROM N.A.
MEDLINE=93186721; PubMed=8444812;
RA Ferrayra R., Sonchizi F., Viale A.M.;
"Cloning, characterization, and functional expression in Escherichia
RT coli of chaperonin (groESL) genes from the phototrophic sulfur
RT bacterium Chromatium vinosum";
RL J. Bacteriol. 175:1514-1523(1993).
RN [2]
RN CHARACTERIZATION.
RX MEDLINE=99008942; PubMed=9790891;
RA Dionisi H.M., Viale A.M.;
" Purification and characterization of chromatium vinosum groEL and
RT groES proteins overexpressed in Escherichia coli cells lacking the
RT endogenous groESL operon.";
RL Protein Expr. Purif. 14:275-282(1998).
CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CC CONDITIONS.
CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
CC 7 SUBUNITS.
CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M99443; AAA23319.1;
DR PIR: B47073; B47073.
DR HSSP: P06139; 1GRL.
DR InterPro: IPR001844;
DR InterPro: IPR002423;
DR Pfam: PF00118; cpn60_TCP1; 1.
DR PRINTS: PRO0298; CHAPERONIN60.
DR PRINTS: PRO0304; TCOMPLEXTCP1.
DR PROSITE: PS00296; CHAPERONIN60; 1.

```

KW Chaperone; ATP-binding.
SQ SEQUENCE 546 AA; 57541 MW; 3250141881C04DD6 CRC64;

Query Match	68.4%;	Score 1852.5;	DB 1;	Length 546;
Best Local Similarity	68.2%;	Pred. No. 8.4e-79;		
Matches 374;	Conservative 70;	Mismatches 101;	Indels 3;	Gaps 2;

[illegible]

	RESULT	8
CH60_RHOCA	CH60_RHOCA	STANDARD; PRT; 545 AA.
P95678:		
U1	01-NOV-1997	(Rel. 35, Created)
DT	01-NOV-1997	(Rel. 35, Last sequence update)
DT	01-NOV-1997	(Rel. 35, Last annotation update)
DE	60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN).	
GN	GROEL.	
GC	Rhodobacter capsulatus (Rhodopseudomonas capsulata).	
OC	Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;	
OC	Rhodobacter.	
OX	NCBI_TaxID=1061;	
RN	[1]	
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.	
RC	STRAIN-B10 / ATCC 33303;	
RX	MEDLINE=96365187; PubMed=8703196;	
RA	Rauber P., Dame G., Sandmeier U., Vandekerckhove J., Beyer P.,	
RA	Tados M.H.;	
RT	Molecular analysis of the Rhodobacter capsulatus chaperonin (groESL).	
RT	operation: purification and characterization of Cpn60.;	
RL	Arch. Microbiol. 166:193-203(1996).	
CC	-I- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND	
CC	PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS	

```

CC      CONDITIONS (BY SIMILARITY).
CC      -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
CC      7 SUBUNITS (BY SIMILARITY).
CC      -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce/
CC      or send an email to license@isb.sib.ch).
CC      -----
CC      DR EMBL; S82593; AAB37532.1; -.
CC      DR HSSP; P06139; IGR1.
CC      DR InterPro; IPR001844; -.
CC      DR InterPro; IPR002423; -.
CC      DR Pfam; PF00118; cpn60_TCPI, 1.
CC      DR PRINTS; PRO0298; CHAPERONIN60.
CC      DR PRINTS; PRO0304; TCOMPLEXTCPI.
CC      DR PROSITE; PS00296; CHAPERONINS_CPN60, 1.
CC      DR Chaperone; ATP-binding.
CC      KW SEQUENCE 545 AA; 57626 MW; 55BCC738DF7DF9A8 CRC64;

```

Query Match	68.1%	Score 1845;	DB 1;	Length 545;
Best Local Similarity	67.2%	Pred. No. 1.9e-78;		
Matches 368; Conservative	76;	Mismatches 100;	Indels 4;	Gaps 2;

```

0Y 1 MASKEILFADAKRFKLSRGVDKLANAKVTLGPRGRNVYIEKSPGSRVITKDCVSVAKEI 60
Db 1 MAKEVEFSTIDARMRKLKGNVILDAKAVTLGPRGRNVYIEKSPGSRVITKDCVSVAKEI 60
0Y 61 ELEDKFENMGQWYKVEVAPRTSDIAGCVTTATVLAQAIYRECVKILVAAGRNMAIKRGI 120
Db 61 ELADKFENMGQWYKVEVASRTRNDEAGCGTTATVLAQAIYRECKMKVAAGRNMDLKRGI 120
0Y 121 DKAVVAATKELSDITKTRPOKEIAOVGTISANSOTTIGNITLIDAMAKGCVITVEEA 180
Db 121 DLATTTVEAIIKAARAPVKDSDVEAOVGTISANGEAQIGRPIADASOKVNEGVIIVEEN 180
0Y 181 KGETTLIDVYEGMKFDRGYSFVPTNPKMCELOMRYTLICNEKRTISKMDLITLEQY 240
Db 181 KGMTEVEYVEGMOFDGTCYISPTFTVPDDMLADLEATYLLHEKTKISSIDPWPVPLEAV 240
0Y 241 AKVNRPLLIIAEDVEGEALATLVVYKRGALQVAVYKAPFGGERRKAMLEDIAILTGGEA 300
Db 241 IQSIRPLIIYAEDVEGEALATLVVYKRGKGLIAAVKAPFGGRRKAMLODIALITLGGY 300
0Y 301 IEDRGITKLENVLSLSIGTAKRVVIDKENTTYVDGACKSEDIKARVYKQIRAOIETSSDY 360
Db 301 ISDOLGAKLENVLTDMIGRAKKVYISKENTTYVDGHDKAEINAVAHNIPTQIEETPSDY 360
0Y 361 DREKLOERLAKLVGAVVIHGAATETEMKEKKDRVEDALNATRAAVEESIVGGTATY 420
Db 361 DREKLOERVAKLVGAVVIRVGGTEVEVERKDRVDALNATRAAVOEBIIYGGVAAV 420
0Y 421 RSIKVLDDIAPADDELAGLNIIRSLSEBPLROIAANAGYEGSIYVEKVEPKD--GQFN 479
Db 421 QAAKKLNDLIGANSDDOAGISIVRRALEAPLROIAENAGYDGVAVVAKGRESADPAGFN 480
0Y 480 AASGEEDLIKAGVIDPKKYTRIALQNAASVLSLLTTECAIEIEKPEKPKDMPGCGMG 539
Db 481 AQTEEVYDMPGFGVYIDPAKYVTRTALDEAASISAGLLITTECMIAEKPEPK---AAPAGMG 537
0Y 540 GMGMDGN 547
Db 538 GMGMDGM 545

```

RESULT	9
CH60_RHILV	
ID	CH60_RHILV
STANDARD;	
PRT;	546 AA

```

AC P34939;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE 60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN).
GN GROEL.
OS Rhizobium leguminosarum (biovar viciae).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=387;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94214663; PubMed=7909257;
RA Wallington E.J., Lund P.A.;
RT "Rhizobium leguminosarum contains multiple chaperonin (cpn60) genes.";
RL Microbiology 140:113-122(1994).
..
-1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CONDITIONS (BY SIMILARITY).
CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
7 SUBUNITS (BY SIMILARITY).
CC -1- INDUCTION: BY HEAT SHOCK.
CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: 120775; AAA26246.1;
DR HSSP: P06139; 1GRL.
DR InterPro: IPR001844;
DR InterPro: IPR002423;
DR Pfam: PF00118; cpn60_TCP1.1.
DR PRINTS: PR00298; CHAPERONIN60.
DR PRINTS: PR00304; TCOMPLEXTCP1.
DR PROSITE: PS00296; CHAPERONIN_CP60.1.
KW Chaperone; ATP-binding; Heat shock.
SQ SEQUENCE 546 AA; 57883 MW; 86C070DEA4B1A8BA CRC64;

Query Match 68.18; Score 1844.5; DB 1; Length 546;
Best Local Similarity 69.88; Pred. No. 2e-78;
Matches 384; Conservative 61; Mismatches 98; Indels 7; Gaps 5;

QY 1 MASKEILFDKAREKLSRGVDKLANAVKTLGPKGRNVIFKSGSPVITKDGVSVAKEI 60
DB 1 MASKEIKRGRTGRREKMLRGVDLADAVKVTLCGPKGRNVILDKSGCAPRLTKDGVSAKEI 60
QY 61 ELFPKFMGAQWKEVAPKTSIDAGDCTTATVLAQAIYREGVYLAAGNPMIAIKIGI 120
DB 61 ELEDKFMGAQWKEVAPKTSIDAGDCTTATVLAQAIYREGVYLAAGNPMIAIKIGI 120
QY 121 DKAVAAVYKKEISDTIKPRDQKEIAOVGTISANSQTTIGNIIAEAMAVGKGVTVEEA 180
DB 121 DLAAADVADKDLQAKAKKISTSEVAOVGTISANGKQGLDIAEMQVGVGEGVITVEEA 180
QY 181 KGLFTLDVEGKMFDRGYSLSYFVTNPEKRVCELDNFIYLCNEKKITSMKDMLEPILEOV 240
DB 181 KTAETLELEVCGMDFDRYLSFYVTNPEKMLADELDFILLHEKKISLNSQMLPYLEAV 240
QY 241 AKVNRPLLITADYVEGELATLVYVKIRGALOAVAVKAPGGERKKAMLEDTAIIITGGEA 300
DB 241 VGTGKPLIIVADEVEGELATLVYVKIRGALOAVAVKAPGGERKKAMLEDTAIIITGTV 300
QY 301 IFEODGILFENVSLSICTAKRVYIDKENTITVCGAGSEDTIKAVNRKIRAOIEETSSDY 360
DB 301 ISEDLGILFESVTLDMLRARAKVSKENTITVDSGAKTIDEGRAVADIKAOIEETISDY 360
QY 361 DREKLOERTLAKLVGVAVIHVGAATETEMKEKKDRVEDALNATRAAVEGIVPGGTAFAV 420

```

```

DB 361 DREKLOERTLAKLVGVAVIRVGSGTEVEVEKEDRIDDALNATRAA0ESIVPGGIALA 420
QY 421 RSIKVLDDIRPADDELAGNIIRSLSEEDPLRQIAANAGYEGSIVKEVREP-KDGFQFN 479
DB 421 RSTRTI-TVKGANDQENGINIVRALSIVRQIAENAGDEASIVGKVLDKNDNGYN 479
QY 480 AASGEYEDLIKAGVIDPKYTRIALQNAASVSLLTTECAIAEKPEPKDMP--MPGG 537
DB 480 AQTSEYEDMLAMGIVDPKLVVTRALQNAASVSLLTTEAMIAEL--PKDAPAGMD-GG 536
QY 538 MGGMGMDGM 547
DB 537 MGGMGMDGM 546

RESULT 10
CH60_BORPE STANDARD; PRT; 547 AA.
ID CH60_BORPE
AC P48210;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN).
GN KOPA OR GROEL OR CPN60.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=TOHAMA / BP338;
RX MEDLINE=95309719; PubMed=7789805;
RA Fernandez R.C., Weiss A.A.;
RT "Cloning and sequencing of the Bordetella pertussis cpn10/cpn60
RT (groEL) homolog."
RL Gene 158:151-152(1995).
..
-1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CONDITIONS (BY SIMILARITY).
CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
7 SUBUNITS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: U12277; AAA74967.1;
DR HSSP: P06139; 1GRL.
DR InterPro: IPR001844;
DR InterPro: IPR002423;
DR Pfam: PF00118; cpn60_TCP1.1.
DR PRINTS: PR00298; CHAPERONIN60.
DR PRINTS: PR00304; TCOMPLEXTCP1.
DR PROSITE: PS00296; CHAPERONIN_CP60.1.
KW Chaperone; ATP-binding.
SQ SEQUENCE 547 AA; 57481 MW; F780D0401F63CF3F CRC64;

Query Match 68.18; Score 1843.5; DB 1; Length 547;
Best Local Similarity 67.48; Pred. No. 2e-78;
Matches 370; Conservative 74; Mismatches 98; Indels 7; Gaps 2;

QY 1 MASKEILFDKAREKLSRGVDKLANAVKTLGPKGRNVIFKSGSPVITKDGVSVAKEI 60
DB 1 MAARKOVLPRADARVAVIGAVVLANAVKTLGPKGRNVIFKSGCAPVITKDGVSVAKEI 60
QY 61 ELEDKFMGAQWKEVAPKTSIDAGDCTTATVLAQAIYREGVYLAAGNPMIAIKIGI 120

```

```

DB 61 EIMDKFENICAOVLKVDYASTSDNACDCTTATVLAQAVVEDELTAKYAAFPNIDLRGI 120
OY 121 DKAVAVTKELSDITKPTROKELIAOVGTISANSPTTIGNIIEAMAKYKGGVITVEEA 180
DB 121 DKAVAAVEBELKLSKPVTTTSKEIAOVGSISSANSASISQIADAMDKYKGGVITVEDG 180
OY 181 KGLETTLDVVEGKPFGRGYLSPFYVFNPEKMCVCLDNPTYLCEKKITTSKMDLPILEOV 240
DB 181 KSIENELDVEEGMQFDGRGILSPFTINSPEKQVALDPPVLTIDKXVSNRDLPLVLEOV 240
OY 241 AKVNRPLLIADVEGEALATLVNKLRCALQVAVAKAPFGERRAMELIDIALITGGEA 300
DB 241 AKSSRPLLIADVEGEALATLVNKLRCALQVAVAKAPFGERRAMELIDIALITGGEA 300
OY 301 IFEDRGIKLENSLSLGTAKRVYIDKENTTYDGAKSDDIARVKQIIRAQIEFTSSDY 360
DB 301 ISEETOMSLKATLLODQOKRTEVAKENTTYIDGAGDGCISIRAVKQIIRAQIEFTSSDY 360
OY 361 DREKLOERLAKLVGVAIVHVGATETEMKEKKDRVEDALNATRAVEEGTIVGGTAFV 420
DB 361 DREKLOERLAKLVGVAIVHVGATETEMKEKKDRVEDALNATRAVEEGTIVGGTAFV 420
OY 421 RSTIKVLDIIRPADDELACINIRSLSEPLRQIAANAGYEGSIVVEKYREPKDGFNA 480
DB 421 RAKQATITGLKGDVADQDNAGIKLLRAVEEPLRTIVTNAGDEASVYVNTVLNGKNGYNA 480
OY 481 ASGEYEDLIKAGVIDPKKVTTRIALONAAVASLTLTTECAIAE----KREPKDMPPGG 536
DB 481 ATEGYDLDVEGQVLDPTKVTYRTALONAAVASLTLTTECAIAE----KREPKDMPPGG 536
OY 537 GMGGMGMD 545
DB 538 GMGGMGMD 546

RESULT 11
CH60_COXBU STANDARD: PRT: 552 AA.
AC P19421:
DT 01-NOV-1990 (rel. 16, Created)
DT 01-NOV-1990 (rel. 16, Last sequence update)
DT 30-NOV-2000 (rel. 39, Last annotation update)
DE 60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN) (HEAT SHOCK PROTEIN B).
GN MOPR OR GROEL OR HTPB.
OS Coxiella burnetii.
OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
OC Coxiella group; Coxiella.
OX NCBI_TaxID=777;
PN [1]

SEQUENCE FROM N.A.
MEDLINE=88139182; PubMed=3343219;
KA Vodka M.H., Williams J.C.:
KT "A heat shock operon in Coxiella burnetii produces a major antigen homologous to a protein in both mycobacteria and Escherichia coli.";
RL J. Bacteriol. 170:1227-1234(1988)
CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS CONDITIONS (BY SIMILARITY).
CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF 7 SUBUNITS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL: M20482; AAA23309.1;

```

```

DR PIR: S39765; S39765.
DR HSRP: P06139; 1GRCL.
DR InterPro: IPR001844; -.
DR InterPro: IPR002423; -.
DR Pfam: PF00118; cpn60_TCP1; 1.
DR PRINTS: PR00298; CHAPERONIN60.
DR PRINTS: PR00304; TCOMPLEXTCP1.
DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
KW Chapterone; ATP-binding; Heat shock.
SQ SEQUENCE 552 AA; 58284 MW; D248857BDB6B1B5E CRC64;

Query Match 68.1%; Score 1843.5; DB 1; Length 552;
Best Local Similarity 66.5%; Pred No. 2.2e-78;
Matches 365; Conservative 82; Mismatches 99; Indels 3; Gaps 3;

OY 1 MASKELIFDPAKREKLSRGVDKLANAVKTLGPKGRNVYKESFGSPVITKQDVSAKEI 60
DB 1 MAKKVLKFSHEVLHAMSRCGEVLANAVKTLGPKGRNVYKESFGAPITTKQDVSAKEI 60
OY 61 ELEDKFEENMGADQVKEVAKRTSDIACDGTITATVLAQAIYREGVKLVAAGRNPMARKGI 120
DB 61 ELEDKFEENMGADQVKEVASTSDAGDGTITATVLAQAIYREGIKAVIAGMNPMDARKGI 120
OY 121 DKAVAVTKELSDITKPTROKELIAOVGTISANSPTTIGNIIEAMAKYKGGVITVEEA 180
DB 121 DKAVTAVAEELKTSKPCQOKIAOVGTISANSPTTIGNIIEAMAKYKGGVITVEEG 180
OY 181 KGLETTLDVVEGKPFGRGYLSPFYVFNPEKMCVCLDNPTYLCEKKITTSKMDLPILEOV 240
DB 181 SGLENLEVEEGMQFDGRGILSPFTINSPEKQVALDPPVLTIDKXVSNRDLPLVLEOV 240
OY 241 AKVNRPLLIADVEGEALATLVNKLRCALQVAVAKAPFGERRAMELIDIALITGGEA 300
DB 241 AKSGRPLLIADVEGEALATLVNKLRCALQVAVAKAPFGERRAMELIDIALITGGEA 300
OY 301 IFEDRGIKLENSLSLGTAKRVYIDKENTTYDGAKSDDIARVKQIIRAQIEFTSSDY 360
DB 301 ISEEVLSLEASLDDLGSKAKRVYTKDPTIIDGSGDAGDKINRVEQIRKELENSSDY 360
OY 361 DREKLOERLAKLVGVAIVHVGATETEMKEKKDRVEDALNATRAVEEGTIVGGTAFV 420
DB 361 DREKLOERLAKLVGVAIVHVGATETEMKEKKDRVEDALNATRAVEEGTIVGGTAFV 420
OY 421 RSTIKVLDIIRPADDELACINIRSLSEPLRQIAANAGYEGSIVVEKYREPKDGFNA 479
DB 421 RVLKSLDSVEENEDQVGEIARAAAPPLSQIVKNTGVQAADVADKLNHRDVMYGYN 480
OY 480 ASGEYEDLIKAGVIDPKKVTTRIALONAAVASLTLTTECAIAEKPCKDMPPGG-M 538
DB 481 AATGEYDMLTENGILDPKVTYRTALONAAVASLTLTTECAIAEKPCKDMPPGG-M 539
OY 539 GGMGMDGM 547
DB 540 GGMGMDGM 548

RESULT 12
CH61_RHOSH STANDARD: PRT: 546 AA.
AC P20110; Q59773;
DT 01-FEB-1991 (rel. 17, Created)
DT 01-NOV-1997 (rel. 35, Last sequence update)
DT 01-NOV-1997 (rel. 35, Last annotation update)
DE 60 KDA CHAPERONIN 1 (PROTEIN CPN60 1) (GROEL PROTEIN 1).
GN Rhodobacter sphaeroides (Rhodospirillum rubrum).
OS Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1063;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HR;

```



```

OY 1 MASKELLFPAKAREKLSRGVDKLANAVKVTGPKGRNVIEKSGSPVITKDGVSVAKEI 60
DB 1 MAEEVFSGSARKKMLVGVNVLADAVKATLGPGRNVIEKSGAPITKDGVSVAKEI 60
OY 61 ELDEKPFEMGQWKEVAPKTSIADGTTATVLAQAIYREGKLVAAAGRNPAIRGDI 120
DB 61 ELKDFEPMGQWKEVAPKTSIADGTTATVLAQAIYREGKLVAAAGRNPAIRGDI 120
OY 121 DKAVAVTKELSDITKTRPOKEIAOVGTTISANSOTTIGNIIEAMAKVKGVIYVEEA 180
DB 121 DKAVAVTKELSDITKTRPOKEIAOVGTTISANSOTTIGNIIEAMAKVKGVIYVEEA 180
OY 181 KGEETTDVVEGKMFDEGYSIPFVTNPEKMWCELDNPIYLCNKKITTSKMDPIIEOVA 240
DB 181 KGEETTDVVEGKMFDEGYSIPFVTNPEKMWCELDNPIYLCNKKITTSKMDPIIEOVA 240
OY 241 AKVNRPLLIADVEGECALATLVNKLQALQVAVKAPGFERKMLIEDIALITGCEA 300
DB 241 AKVNRPLLIADVEGECALATLVNKLQALQVAVKAPGFERKMLIEDIALITGCEA 300
OY 301 IFEDRGKLENVSLSSIGTKRRVYIDKENTTYDVGAKSEDIKARVQIRAOIETSSDY 360
DB 301 IFEDRGKLENVSLSSIGTKRRVYIDKENTTYDVGAKSEDIKARVQIRAOIETSSDY 360
OY 361 DREKIOERLAKLVGVAVIHGATETEMKEKKDRVEDALNATRAAEEGIVPGCGTAFAV 420
DB 361 DREKIOERLAKLVGVAVIHGATETEMKEKKDRVEDALNATRAAEEGIVPGCGTAFAV 420
OY 421 RSIKLVDDIKPADDELALNITIRSLPEPLROIAANAGYEGSIYVEKREPKDGFEN 480
DB 421 RSIKLVDDIKPADDELALNITIRSLPEPLROIAANAGYEGSIYVEKREPKDGFEN 480
OY 481 ASGEYEDLIKAGYIDPKKVTIRLQNAASVALLITTECAIAEKPEPKDMPMP 539
DB 481 ASGEYEDLIKAGYIDPKKVTIRLQNAASVALLITTECAIAEKPEPKDMPMP 539
OY 540 GNGMGDM 547
DB 540 GNGMGDM 547
OY 541 GNGMGDM 548
DB 541 GNGMGDM 548

```

RESULT 14
CH61_BRAJA STANDARD: PRT: 539 AA.
AC P77829;
DT 01-NOV-1997 (rel. 35, Created)
DT 01-NOV-1997 (rel. 35, Last sequence update)
DE 60 KDA CHAPERONIN 1 (PROTEIN CPN60 1) (GROEL PROTEIN 1).
GN GROEL
Bradyrhizobium japonicum.
Bacteria: Proteobacteria; alpha subdivision; Rhizobiaceae group;
Bradyrhizobium group; Bradyrhizobium.
NCBI_TaxID=375;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96404779; PubMed=8808920;
RA Harberhaus F., Weiglhofer W., Fischer H.M., Hennecke H.;
RT "The Bradyrhizobium japonicum (pbn) gene encoding a sigma 32-like
protein is part of a unique heat shock gene cluster together with
groEL and three small heat shock genes.";
RL J. Bacteriol. 178:5337-5346(1996).
CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CC CONDITIONS (BY SIMILARITY).
CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
CC 7 SUBUNITS (BY SIMILARITY).
CC -1- INDUCTION: INDUCED BY HEAT SHOCK.
CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U55047; AAC44753.1; -
DR HSSP: P06139; 1GRL.
DR InterPro: IPR001844; -
DR InterPro: IPR002423; -
DR Pfam: PF00118; cpn60_TCP1.1.
DR PRINTS: PR00298; CHAPERONIN60.
DR PRINTS: PR00304; TCOMPLEXCP1.
DR PROSITE: PS00296; CHAPERONINS_CP60.1.
DR KAPR: KAPR00000; CHAPERONIN; Multigene family; Heat shock.
FT INIT_MET 0
FT SIMILARITY 0
SQ SEQUENCE 539 AA; 57419 MW; CE9641A0FB1B54 CRC64;

Query Match 67.7%; Score 1833.5; DB 1; Length 539;
Best Local Similarity 69.5%; Pred. No. 6; 2e-78;
Matches 371; Conservative 68; Mismatches 92; Indels 3; Gaps 2;

OY 2 ASKEILFPAKAREKLSRGVDKLANAVKVTGPKGRNVIEKSGSPVITKDGVSVAKEI 61
DB 1 AAEVVFSTDAKRVFLKGVDTLQANAVKVTGPKGRNVIEKSGAPITKDGVSVAKEI 60
OY 62 LEDKPFEMGQWKEVAPKTSIADGTTATVLAQAIYREGKLVAAAGRNPAIRGDI 121
DB 62 LEDKPFEMGQWKEVAPKTSIADGTTATVLAQAIYREGKLVAAAGRNPAIRGDI 120
OY 122 KAVVAATKELSDITKTRPOKEIAOVGTTISANSOTTIGNIIEAMAKVKGVIYVEEA 181
DB 122 KAVVAATKELSDITKTRPOKEIAOVGTTISANSOTTIGNIIEAMAKVKGVIYVEEA 180
OY 182 GLETTTDVVEGKMFDEGYSIPFVTNPEKMWCELDNPIYLCNKKITTSKMDPIIEOVA 241
DB 182 GLETTTDVVEGKMFDEGYSIPFVTNPEKMWCELDNPIYLCNKKITTSKMDPIIEOVA 240
OY 242 KVNRPPLIADVEGECALATLVNKLQALQVAVKAPGFERKMLIEDIALITGCEA 301
DB 242 KVNRPPLIADVEGECALATLVNKLQALQVAVKAPGFERKMLIEDIALITGCEA 300
OY 302 IFEDRGKLENVSLSSIGTKRRVYIDKENTTYDVGAKSEDIKARVQIRAOIETSSDY 361
DB 302 IFEDRGKLENVSLSSIGTKRRVYIDKENTTYDVGAKSEDIKARVQIRAOIETSSDY 360
OY 361 DREKIOERLAKLVGVAVIHGATETEMKEKKDRVEDALNATRAAEEGIVPGCGTAFAV 421
DB 361 DREKIOERLAKLVGVAVIHGATETEMKEKKDRVEDALNATRAAEEGIVPGCGTAFAV 420
OY 421 RSIKLVDDIKPADDELALNITIRSLPEPLROIAANAGYEGSIYVEKREPKDGFEN 480
DB 421 RSIKLVDDIKPADDELALNITIRSLPEPLROIAANAGYEGSIYVEKREPKDGFEN 480
OY 481 ASGEYEDLIKAGYIDPKKVTIRLQNAASVALLITTECAIAEKPEPKDMPMP 534
DB 481 ASGEYEDLIKAGYIDPKKVTIRLQNAASVALLITTECAIAEKPEPKDMPMP 532

```

RESULT 15
CH60_AGRU STANDARD: PRT: 544 AA.
AC P30779;
DT 01-JUL-1993 (rel. 26, Created)
DT 01-JUL-1993 (rel. 26, Last sequence update)
DT 30-MAY-2000 (rel. 39, Last annotation update)
DE 60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN).
GN MOPA OR GROEL.
OS Agrobacterium tumefaciens.
OC Bacteria: Proteobacteria; alpha subdivision; Rhizobiaceae group;
Rhizobiaceae; Agrobacterium.
NCBI_TaxID=362;

RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-C58;
 RX MEDLINE=93259955; PubMed=8098329;
 RA Segal G., Ron E.Z.;
 RT "Heat shock transcription of the groESL operon of Agrobacterium
 RL J. Bacteriol. 175:3083-3088(1993).
 CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
 CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
 CC CONDITIONS (BY SIMILARITY).
 CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS, OF
 CC 7 SUBUNITS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC -----
 CC EMBL: X68263; CAA48331.1;
 CC DR PIR: S23918; S23918.
 CC DR HSSP: P06139; 1GRL.
 CC DR InterPro: IPR001844;
 CC DR InterPro: IPR002423;
 CC DR Pfam: PF00118; Cpn60_TCP1; 1.
 CC DR PRINTS: PR00298; CHAPERONIN60.
 CC DR PRINTS: PR00304; TCOMPLEXTCP1.
 CC DR PROSITE: PS00296; CHAPERONIN5_CPN60; 1.
 CC KW Chaperone; ATP-binding.
 CC SEQUENCE: 544 AA; 57547 MW; 1E251B3F7BE2E22F CRC64;

Query Match 67.5%; Score 1828.5; DB 1; Length 544;
 Best Local Similarity 68.9%; Pred. No. 1,1e-77;
 Matches 374; Conservative 70; Mismatches 96; Indels 3; Gaps 3;

QY 1 MASKEILFDKAREKLSGVNDKLANAVKVLGPGGRNVYIEKSPGSPVITKDGVSVAKEI 60
 DB 1 MAKEVFPKASAREKMKLGVDILADVAVTLGPKGRNVYIDKSGAPITKDGVSVAKEI 60
 QY 61 ELEDKPFEMGAOMVKFVAPKTSAGDSTTATVLAQAIYREGVKLVAGRNPAIKRGI 120
 DB 61 ELEDKPFEMGAOMVKFVAPKTSAGDSTTATVLAQAIYREGVKLVAGRNPAIKRGI 120
 QY 121 DKAVVAATKELSDITKPTROKELIAOVGTISANSDDTIGNIIAEAMAKVKGKGVITVEEA 180
 DB 121 DLAAVAEVKDLQAKAKKINTISEEVAQVOTISANGEROIGLDIAEMQVNGEGVITVEEA 180
 QY 181 KGLFTYLDVVEGKMFDRGVYSPYFVTPNEKMWCELDNFIYCNEKKITSMKMDLPILEOV 240
 DB 181 KTAETELEVVEGMOFDRGVYSPYFVTPNEKMWCELDNFIYLHEKKISLNQAMLPYLEAV 240
 QY 241 AKVNRPLIIIAEDVEGELATLVYVNLKGLAQVAVKAPGGERKKAMLEDIAIITGGEA 300
 DB 241 VOTCKPLVLLIAEDVEGELATLVYVNLKGLAQVAVKAPGGERKKAMLEDIAIITGCTV 300
 QY 301 IFEDRGIKLENVSLSTGAKRVVIDKENTIVDGAGKSDIKARVKOIRAOIETSSDY 360
 DB 301 ISEDLGKIKLESVTLDMCKSKKVSISKNTTIVDGAGOKSDIEGRVADIKAOIETTSY 360
 QY 361 DREKLOERLAKLVGVAVIHVGAATEEMKEKKDRVEDALNATRAAVEGIVPGGTAFAV 420
 DB 361 DREKLOERLAKLVGVAVIRVGGSTEVEVEKEDRIDDALNATRAAVEGIVPGGVAL 420
 QY 421 RSTIVLDDIKPADDELACNLTIRSLIEPLRQIAANAGYEGSTIVEKVEEP-KDGFQFN 479
 DB 421 RSTIKI-TVKGVNDQDEGINIVKALOSLVRQIAENAGDEASTIVGKILDKNEDNYGN 479
 QY 480 AASGEYEDLIKAGVIDPKKVTRIALQNAASVASILLTTECAIAEKPEPKKDMF-MPGGM 538

Db 480 AQTGEYDILATGIVDPKVVVTRALQNAASVASVLTTEAMIAELPKKESAMPQMPGGM 539
 QY 539 GGM 541
 Db 540 GGM 542

Search completed: August 1, 2001, 17:26:34
 Job time: 470 sec

Thu Aug 2 08:05:08 2001

us-09-077-574a-2.rsp

Page 12

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2001, 17:18:14 ; Search time 59 seconds

(without alignments)
1228.867 Million cell updates/sec

Title: US-09-077-574a-2
Sequence: 1 MASKEDKAKREKLSRGV.....KDMPMGGGGMGMDGW 548

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

arched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SPREMBL_16:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.unclassified:*
13: sp.vertebrate:*
14: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2708	100.0	548	2	087888 lawsonia in
2	1868.5	69.0	546	2	09F712 burkholderi
3	1862.5	68.8	547	2	09PFP2 xylella fas
4	1852.5	68.4	540	2	09XCA9 rhodothermu
5	1834	67.7	542	2	09LAP7 rhizobium 1
6	1804.5	66.6	547	2	09XAU7 alteromonas
7	1768	65.3	539	2	09RC20 bacillus sp
8	1768	65.3	546	2	031198 leptostrila
9	1761.5	65.0	544	2	09KMR7 vibrio chol
10	1755.5	64.8	544	2	09L690 rhizobium 1
11	1754	64.8	539	2	09E2V4 bacillus st
12	1753	64.3	548	2	09L7P5 vibrio para
13	1741	64.3	540	2	09XW14 methylovoru
14	1736	64.1	538	2	033688 sitophilus
15	1722	63.6	544	2	09XG73 bacteroides
16	1720	63.5	543	2	09RA44 thermus aqu
17	1708.5	63.1	548	2	09RW09 deinococcus
18	1706	63.0	555	2	09X603 primary end
19	1704.5	62.9	550	2	09S303 rickettsia

20	1672.5	61.8	548	2	008499
21	1671.5	61.7	548	2	008500
22	1663.5	61.4	540	2	09X455
23	1660	61.3	503	2	087271
24	1640	60.6	497	2	09X4G5
25	1637.5	60.5	544	2	032847
26	1637	60.5	530	2	09KLC6
27	1635	60.4	538	2	09WYX6
28	1634.5	60.4	538	2	09E2V1
29	1630	60.2	541	2	09KX05
30	1626	60.0	497	2	09S6B0
31	1616	59.7	497	2	09X4G4
32	1614.5	59.6	544	2	09K2U4
33	1595.5	58.9	540	2	09K1V7
34	1594.5	58.9	483	2	087272
35	1593.5	58.8	540	2	09KRF0
36	1588.5	58.7	541	2	09KGM1
37	1588	58.6	543	2	09KJ23
38	1583.5	58.5	540	2	068324
39	1563.5	57.7	533	2	09FDS2
40	1560.5	57.6	539	2	097086
41	1552	57.3	483	2	087270
42	1524.5	56.3	463	2	087274
43	1515.5	56.0	474	2	087269
44	1475.5	54.5	594	3	094110
45	1473.5	54.4	575	5	09Y108

ALIGNMENTS

RESULT 1
ID 087888 PRELIMINARY; PRT; 548 AA.
AC 087888;
DT 01-NOV-1998 (TEMBLrel. 08, Created)
DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
DE 60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN).
OS Lawsonia intracellularis.
OC Plasmid pISI-2.
OC Bacteria; Proteobacteria; delta subdivision; Lawsonia <bacteria>.
OX NCBI_TaxID:29546;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:9836497; PubMed:9720028;
RA Dale C.J.H., Moses E.R., Ong C.C., Morrow C.J., Reed M.B., Hasse D., Struelens R.A.;
RT "Identification and sequencing of the gro operon and flanking genes of Lawsonia intracellularis: use in phylogeny."
RL Microbiology 144:0-0(0).
CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS CONDITIONS (BY SIMILARITY).
CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF 7 SUBUNITS (BY SIMILARITY).
CC -1- SIMILARITY: TO THE CHAPERONIN CPN60 (60KD SUBUNIT) DOMAIN.
CC EMBL: U45241; AAC36500.1; -;
DR HSSP; P06139; 1GRL.
DR InterPro: IPR001844; -;
DR InterPro: IPR002423; -;
DR Pfam: PF00118; cpn60_Tcpl. 1.
DR PRINTS; PR00298; CHAPERONIN60.
DR PROSITE; PR00304; TCOMPLEXTCPL.
DR PROSITE; PS00296; CHAPERONIN_CP60; 1.
KW ATP-binding; Chaperone; Plasmid.
SQ SEQUENCE 548 AA; 58605 MW; 6388C431E63EA98 CRC64;

Query Match 100.0%; Score 2708; DB 2; Length 548;
Best Local Similarity 100.0%; Pred. No. 1.4e-119;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MASKEILFDPAKAREKLSRGVDKLANAVKVTLGPKGRNVVIEKSFSGPVITTKDGVSAKEI 60
DB 1 MASKEILFDPAKAREKLSRGVDKLANAVKVTLGPKGRNVVIEKSFSGPVITTKDGVSAKEI 60
OY 61 ELEDKFEENNGAOMKEVAPKRTSDIACDGTATVLAQAALYREGVKLVAAGRNPMALIKRGI 120
DB 61 ELEDKFEENNGAOMKEVAPKRTSDIACDGTATVLAQAALYREGVKLVAAGRNPMALIKRGI 120
OY 121 DKAAVAATVKELSDITKPTFDOKELIAOVGTISANSOTTIGNITAEAMAKKGGVITVEEA 180
DB 121 DKAAVAATVKELSDITKPTFDOKELIAOVGTISANSOTTIGNITAEAMAKKGGVITVEEA 180
OY 181 KGLLETTLDVVEGKMPDRGLSPFTNPKMVCEDLNPYLILNEKITSKMDPLILEOV 240
DB 181 KGLLETTLDVVEGKMPDRGLSPFTNPKMVCEDLNPYLILNEKITSKMDPLILEOV 240
OY 241 AKVNRPLLIADVEGEALATLVNKLRCALOVAAKAPFGERRKAMLEDAIALLTGGEA 300
DB 241 AKVNRPLLIADVEGEALATLVNKLRCALOVAAKAPFGERRKAMLEDAIALLTGGEA 300
OY 301 IFEDNGIKLENNVSLSLGAKRVIDKENTTTVDGAKSEDIKARYKOIRAOIEETSSDY 360
DB 301 IFEDNGIKLENNVSLSLGAKRVIDKENTTTVDGAKSEDIKARYKOIRAOIEETSSDY 360
OY 361 DREKLOERLAKLVGVAIVHGAATETEMKEKKDRVEDALNMTRAVEGIVPGGTAIV 420
DB 361 DREKLOERLAKLVGVAIVHGAATETEMKEKKDRVEDALNMTRAVEGIVPGGTAIV 420
OY 421 RSIKVLDIKPADDELACGLNIRSLLEPLRQIAANAGYEGSIYVEKREPKDGFNA 480
DB 421 RSIKVLDIKPADDELACGLNIRSLLEPLRQIAANAGYEGSIYVEKREPKDGFNA 480
OY 481 ASGEVEDLIKAGVIDPKKVTTRIALONASVASLLTTGCAIAEKPEPKDMPPGGMKG 540
DB 481 ASGEVEDLIKAGVIDPKKVTTRIALONASVASLLTTGCAIAEKPEPKDMPPGGMKG 540
OY 541 MCGMDGM 548
DB 541 MCGMDGM 548

```

RESULT 2

09F712 PRELIMINARY: PRT: 546 AA.

AC 09F712: 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE CHAPERONIN GROEL.

OS Burkholderia pseudomallei (Pseudomonas pseudomallei).

CC Burkholderia; Proteobacteria; beta subdivision; Burkholderia group;

NCBI_TaxID=28450;

RP SEQUENCE FROM N.A.

RA Yuen, Woo P.C.Y., Leung P.K.L.:

RT "Burkholderia pseudomallei groEL gene."

RL Submitted (Jul-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF287633; AAC32927.1.

SO SEQUENCE 546 AA: 57145 MW: E9E9366EB8BD6ABD CRC64:

Query Match 69.0%; Score 1868.5; DB 2; Length 546;

Best Local Similarity 68.2%; Pred. No. 3.2e-80;

Matches 373; Conservative 75; Mismatches 96; Indels 3; Gaps 2;

```

OY 1 MASKEILFDPAKAREKLSRGVDKLANAVKVTLGPKGRNVVIEKSFSGPVITTKDGVSAKEI 60
DB 1 MASKEILFDPAKAREKLSRGVDKLANAVKVTLGPKGRNVVIEKSFSGPVITTKDGVSAKEI 60
OY 61 ELEDKFEENNGAOMKEVAPKRTSDIACDGTATVLAQAALYREGVKLVAAGRNPMALIKRGI 120
DB 61 ELEDKFEENNGAOMKEVAPKRTSDIACDGTATVLAQAALYREGVKLVAAGRNPMALIKRGI 120
OY 121 DKAAVAATVKELSDITKPTFDOKELIAOVGTISANSOTTIGNITAEAMAKKGGVITVEEA 180
DB 121 DKAAVAATVKELSDITKPTFDOKELIAOVGTISANSOTTIGNITAEAMAKKGGVITVEEA 180
OY 181 KGLLETTLDVVEGKMPDRGLSPFTNPKMVCEDLNPYLILNEKITSKMDPLILEOV 240
DB 181 KGLLETTLDVVEGKMPDRGLSPFTNPKMVCEDLNPYLILNEKITSKMDPLILEOV 240
OY 241 AKVNRPLLIADVEGEALATLVNKLRCALOVAAKAPFGERRKAMLEDAIALLTGGEA 300
DB 241 AKVNRPLLIADVEGEALATLVNKLRCALOVAAKAPFGERRKAMLEDAIALLTGGEA 300
OY 301 IFEDNGIKLENNVSLSLGAKRVIDKENTTTVDGAKSEDIKARYKOIRAOIEETSSDY 360
DB 301 IFEDNGIKLENNVSLSLGAKRVIDKENTTTVDGAKSEDIKARYKOIRAOIEETSSDY 360
OY 361 DREKLOERLAKLVGVAIVHGAATETEMKEKKDRVEDALNMTRAVEGIVPGGTAIV 420
DB 361 DREKLOERLAKLVGVAIVHGAATETEMKEKKDRVEDALNMTRAVEGIVPGGTAIV 420
OY 421 RSIKVLDIKPADDELACGLNIRSLLEPLRQIAANAGYEGSIYVEKREPKDGFNA 480
DB 421 RSIKVLDIKPADDELACGLNIRSLLEPLRQIAANAGYEGSIYVEKREPKDGFNA 480
OY 481 ASGEVEDLIKAGVIDPKKVTTRIALONASVASLLTTGCAIAEKPEPKDMPPGGMKG 540
DB 481 ASGEVEDLIKAGVIDPKKVTTRIALONASVASLLTTGCAIAEKPEPKDMPPGGMKG 540
OY 541 MCGMDGM 547
DB 539 MCGM-GM 544

```

RESULT 3

09F72 PRELIMINARY: PRT: 547 AA.

AC 09F72: 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE 60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN).

GN XP0615.

OS Xylella fastidiosa.

CC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;

OC Xylella.

OX NCBI_TaxID=2371;

RP SEQUENCE FROM N.A.

RC STRAIN=9A5C;

RX MEDLINE=20365717; PubMed=10910347;

RA Simpson A.J.G., Relnach F.C., Artuda P., Abreu F.A., Acencio M.,

RA Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,

RA Barros M.H., Bonaccorsi E.D., Bordin L.S., Bove J.M., Bionnes M.R.S.,

RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carret H.,

RA Coutinho L.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,

RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorty H.,

RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,

RA Fria J.S., Franco S.C., Franco M.C., Frohne M., Furian L.R.,

RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,

RA Ho P.L., Hohnselt J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,

RA Krieger J.E., Kuramae E.E., Laloret F., Lambais M.R., Leite L.C.C.,

RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,

RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,

RA Marques M.V., Martins E.A.L., Martins S.L., Matsukuma A.Y.,

RA Menck C.F.M., Miyaca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,

RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,

RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,

RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,

RA Peixoto B.R., Pereira G.A.G., Pereira H.A.Jr., Pesquero J.B.,

RA Queglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,

RA de Rosa V.E.Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,

RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A.Jr

RA da silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Trufl D., Tsai S.M., Tsunako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.,
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*,"
 RL Nature 406:151-159(2000).
 CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
 CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
 CC CONDITIONS (BY SIMILARITY).
 CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
 CC 7 SUBUNITS (BY SIMILARITY).
 CC -1- SIMILARITY: TO THE CHAPERONIN CPN60 (60KD SUBUNIT) DOMAIN.
 DR EMBL: AE00307; AAF83425.1;
 DR InterPro: IPR001844;
 DR InterPro: IPR002423;
 DR Pfam: PF00118; cpn60_TCP1; 1.
 DR PRINTS: PR00296; CHAPERONIN60.
 DR PRINTS: PR00304; TCOMPLEXTCP1.
 DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
 KM ATP-binding; Chaperone.
 SO SEQUENCE 547 AA; 57757 MW; 30BE7F937CA7A9D8 CRC64;

Query Match 68.8%; Score 1862.5; DB 2; Length 547;
 Best Local Similarity 66.8%; Pred. No. 6.1e-80;
 Matches 366; Conservative 80; Mismatches 97; Indels 5; Gaps 2;

OY 1 MASKELLFDAKAREKLSRGVDKLANAVKVLGPKGRNVILEKSGSVITKKGVSVAKEI 60
 DB 1 MAKEIIFSEKARSRMVHGNLANAVKATLGPKGRNVILEKSGSVITKKGVSVAKEI 60
 OY 61 ELDKFEENMGAWKVEAPRTSDIAGDGTATVLAQAIAREGKVLAAAGNPMAIRGI 120
 DB 61 ELADKFEENMGAWKVEAPRTSDIAGDGTATVLAQAIAREGKVLAAAGNPMAIRGI 120
 OY 121 DKAVVAATKELSDITKPTROKKEIAQVGTISANSDDTIGNIIAEAMAKVKGVIYEEA 180
 DB 121 DKAVVAATKELSDITKPTROKKEIAQVGTISANSDDTIGNIIAEAMAKVKGVIYEEA 180
 OY 181 KGETTLIDVVEGKMFDRGYLSPYFVTPPEKMYCELDNPYLLCNEKKTSMKMLPILEOV 240
 DB 181 KGETTLIDVVEGKMFDRGYLSPYFVTPPEKMYCELDNPYLLCNEKKTSMKMLPILEOV 240
 OY 241 AKVNRPLIIAEDVEGALATLVNKLRGALOVAVAPGGERKRAMLEDIALITGGEA 300
 DB 241 AKVNRPLIIAEDVEGALATLVNKLRGALOVAVAPGGERKRAMLEDIALITGGEA 300
 OY 301 IFEDRGIKLENVSLSGTAKRNVIDKENTTIYDGAKSSEDIKARVQIRAOIETSSDY 360
 DB 301 ISEEVGSLKATTSIHGKAKKVRVSKENTTIIDGIGDVAINGKRVQIRAOIETSSDY 360
 OY 361 DREKLOERLAKLVGGVAVIHVGAATEEMKEKKORVEDALNATRAAVEEGIVPGGTAIV 420
 DB 361 DREKLOERLAKLVGGVAVIHVGAATEEMKEKKORVEDALNATRAAVEEGIVPGGTAIV 420
 OY 421 RSTIVLDDIKPADDELALGNIIRSLSEPLROIANAAGVGSIVKVRPKGPGFNA 480
 DB 421 RSTIVLDDIKPADDELALGNIIRSLSEPLROIANAAGVGSIVKVRPKGPGFNA 480
 OY 481 ASGEEDLIRAGVIDPKVTTRIALONAAVASALLTTECAIAEKPEPKDMPMP---GGG 537
 DB 481 ATGFGGMVNLGILDPKVTTRIALONAAVASALLTTECAIAEKPEPKDMPMP---GGG 537
 OY 538 MCGMGMD 545
 DB 539 MCGMGMD 546

RESULT 4
 O9XCA9 PRELIMINARY; PRT; 540 AA.
 AC O9XCA9;
 DT 01-NOV-1999 (Tremblrel. 12, Created)

DT 01-NOV-1999 (Tremblrel. 12, last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, last annotation update)
 DE 60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN).
 GN GROEL.
 OS Rhodothermus marinus.
 OC Bacteria: CFB group; Rhodothermus group; Rhodothermus.
 OX NCBI_TaxID=29549;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ITI 376.
 RA Thoroldsdottir E.T.T., Backman V.M., Blondal T.,
 RA Thorjarnardottir S.H., Palsdottir A., Hauksdottir H.,
 RA Kristjansdottir S., Eggertsson G.;
 RT "Heat shock in Rhodothermus marinus: Cloning and sequence analysis of
 RT the groEL, dnaK and dnaJ genes";
 RL Submitted (Apr-1989) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
 CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
 CC CONDITIONS (BY SIMILARITY).
 CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
 CC 7 SUBUNITS (BY SIMILARITY).
 CC -1- SIMILARITY: TO THE CHAPERONIN CPN60 (60KD SUBUNIT) DOMAIN.
 DR EMBL: AF145252; AAD37976.1;
 DR HSSP: P06139; IGRU.
 DR InterPro: IPR001844;
 DR InterPro: IPR001844;
 DR InterPro: IPR002423;
 DR Pfam: PF00118; cpn60_TCP1; 1.
 DR PRINTS: PR00296; CHAPERONIN60.
 DR PRINTS: PR00304; TCOMPLEXTCP1.
 DR ProDom: PD000930; 1.
 DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
 KM ATP-binding; Chaperone; Heat shock.
 SO SEQUENCE 540 AA; 57666 MW; 6867448B7BEF18EC CRC64;

Query Match 68.2%; Score 1852.5; DB 2; Length 540;
 Best Local Similarity 69.2%; Pred. No. 1.8e-79;
 Matches 371; Conservative 71; Mismatches 93; Indels 1; Gaps 1;

OY 1 MASKELLFDAKAREKLSRGVDKLANAVKVLGPKGRNVILEKSGSVITKKGVSVAKEI 60
 DB 1 MAKEIIFNADARMLKRGVDKLANAVKVLGPKGRNVILEKSGSVITKKGVSVAKEI 60
 OY 61 ELDKFEENMGAWKVEAPRTSDIAGDGTATVLAQAIAREGKVLAAAGNPMAIRGI 120
 DB 61 ELDKFEENMGAWKVEAPRTSDIAGDGTATVLAQAIAREGKVLAAAGNPMAIRGI 120
 OY 121 DKAVVAATKELSDITKPTROKKEIAQVGTISANSDDTIGNIIAEAMAKVKGVIYEEA 180
 DB 121 DKAVVAATKELSDITKPTROKKEIAQVGTISANSDDTIGNIIAEAMAKVKGVIYEEA 180
 OY 181 KGETTLIDVVEGKMFDRGYLSPYFVTPPEKMYCELDNPYLLCNEKKTSMKMLPILEOV 240
 DB 181 KGETTLIDVVEGKMFDRGYLSPYFVTPPEKMYCELDNPYLLCNEKKTSMKMLPILEOV 240
 OY 241 AKVNRPLIIAEDVEGALATLVNKLRGALOVAVAPGGERKRAMLEDIALITGGEA 300
 DB 241 VQGRPLIIAEDVEGALATLVNKLRGALOVAVAPGGERKRAMLEDIALITGGEA 300
 OY 301 IFEDRGIKLENVSLSGTAKRNVIDKENTTIYDGAKSSEDIKARVQIRAOIETSSDY 360
 DB 301 ISEBKGRLENATLDYIGQAEIRIVDKDNTTIYDGAKSDPAIKRANQIRAOIETSSDY 360
 OY 361 DREKLOERLAKLVGGVAVIHVGAATEEMKEKKORVEDALNATRAAVEEGIVPGGTAIV 420
 DB 361 DREKLOERLAKLVGGVAVIHVGAATEEMKEKKORVEDALNATRAAVEEGIVPGGTAIV 420
 OY 421 RSTIVLDDIKPADDELALGNIIRSLSEPLROIANAAGVGSIVKVRPKGPGFNA 480
 DB 421 RSTIVLDDIKPADDELALGNIIRSLSEPLROIANAAGVGSIVKVRPKGPGFNA 480
 OY 481 ASGEEDLIRAGVIDPKVTTRIALONAAVASALLTTECAIAEKPEPKDMPMPG 535

DB 481 QREFCNLLGCVIDPTKVAITALENAASVAGILLTTEAVVAEKEPEKKAAPSPG 536

RESULT 5

091691 PRELIMINARY: PRT: 542 AA.

AC 091691;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE 60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN).
GN CPN60-2.
OS Rhizobium leguminosarum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=384;
RN 11
RP SEQUENCE FROM N.A.
RA Rodriguez-Quinones F., Lund P.A.;
RT "Sequence of the cpn2 operon of Rhizobium leguminosarum."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
-1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CONDITIONS (BY SIMILARITY).
CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
7 SUBUNITS (BY SIMILARITY).
CC -1- SIMILARITY: TO THE CHAPERONIN CPN60 (60KD SUBUNIT) DOMAIN.
DR EMBL: AF239163; AAF64160.1; -
DR InterPro: IPR001844; -
DR InterPro: IPR002423; -
DR InterPro: IPR002633; -
DR Pfam: PF00118; cpn60_TCP1.1;
DR PRINTS: PR00298; CHAPERONIN60.
DR PRODOM: PD004452; -1-
DR PRINTS: PD00304; TCOMPLEXTCP1.
DR PROSITE: PS00296; CHAPERONIN_CP60; 1.
KM ATP-binding; Chaperone.
SQ SEQUENCE 542 AA: 57515 MW: 443645703ADF2187 CAC64:

Query Match 67.7%; Score 1834; DB 2; Length 542;
Best Local Similarity 68.3%; Pred. No. 1.3e-78;
Matches 373; Conservative 66; Mismatches 101; Indels 6; Gaps 2;

OY 1 MASKELLDARKREKLSGVCKLANAVVYTGPKGRNVIEKSFSPVYTKGVSVAKEI 60
DB 1 MAKEIKFSTEARERKMLRGVDILANAVKATLCPKGRNVIDKSFAPRTTKGVSVAKEI 60
OY 61 ELEDKFENNAGOMVEKVAAPKTSIDAGDGTATVLAQAIYRGVVLVAGRPMAIKRGI 120
DB 61 ELEDKFENNAGOMVEKVAAPKTSIDAGDGTATVLAQAIYRGVVLVAGRPMAIKRGI 120
OY 121 DKAVAVAVTKELSDITRKPRDOKEIAOVGTISANSOTTTIGNITIAEMAKVKGCVITVEEA 180
DB 121 DLAVAIAVAELKANKARKISNNSEIAOVGTISANGDAEIGRFLEAMEKNGDGVITVEEA 180
OY 181 KLEETTLVVVEGMRKFDKGLSPFYTNPEKMCVCLDNPIILCEKTKISMKMLPILVEY 240
DB 181 KALETELEVEGMRKFDKGLSPFYTNPEKMCVCLDNPIILCEKTKISMKMLPILVEY 240
OY 241 AKVNPRLIIADVEGEALATLVVVKLRGALQOVAVKAPGCGERRKAMLEDIAITGCGA 300
DB 241 VQSSNPRLIIADVEGEALATLVVVKLRGALQOVAVKAPGCGERRKAMLEDIAITGCGA 300
OY 301 IFEEDGRIKLENVSLSLCTAKRNVIDKENTTIVDGAGSEDIKARVKQIRAOIEETSSDY 360
DB 301 ISEDGRIKLENVSLSLCTAKRNVIDKENTTIVDGAGSEDIKARVKQIRAOIEETSSDY 360
OY 361 DREKIOERLAKLVGVAVVHGAATEKKEKDRVEDALNTRAAGEGIVPGGATFV 420
DB 361 DREKIOERLAKLVGVAVVHGAATEKKEKDRVEDALNTRAAGEGIVPGGATFV 420

OY 421 RSIVKLDQIKPADDELANTIRSLERLEPLROIANAGEGSIVYKVEPKD-CGCPN 479
DB 421 RAVKALDAVKTANGDQVVDIVRRVVEAPARQIEMNAGESVYKRLRESSESYGN 480
OY 480 AASGEYEDLIKAGVIDPKKVTIRIALONASVASLLTTECAIAEKPEPKDPMGCGG 539
DB 481 AOTGEYGLYACGVIDPAVAVVTRALQDASISGLVTTTAMIAEKPKKAPPPMAG--- 537
OY 540 GKGCHD 545
DB 538 --PGMD 541

RESULT 6

09XAUT7 PRELIMINARY: PRT: 547 AA.

AC 09XAUT7;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE 60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN).
GN GROEL.
OS Alteromonas haloplanktis.
OC Bacteria; Proteobacteria; gamma subdivision; Alteromonadaceae;
OC Pseudalteromonas.
OX NCBI_TaxID=228;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=TAC125;
RA Tosco A., Birolo L., Scaloni A., Sanna G., Martino G.;
RT "A GroEL-like protein from the psychrotrophic bacterium
Pseudalteromonas haloplanktis TAC125."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
-1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CONDITIONS (BY SIMILARITY).
CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
7 SUBUNITS (BY SIMILARITY).
CC -1- SIMILARITY: TO THE CHAPERONIN CPN60 (60KD SUBUNIT) DOMAIN.
DR EMBL: AJ243594; CAB50775.1; -
DR HSP: P06139; 1GRU.
DR InterPro: IPR001844; -
DR InterPro: IPR002423; -
DR Pfam: PF00118; cpn60_TCP1.1;
DR PRINTS: PR00298; CHAPERONIN60.
DR PROSITE: PS00296; CHAPERONIN_CP60; 1.
KM ATP-binding; Chaperone.
SQ SEQUENCE 547 AA: 57143 MW: P82420C58673B338 CAC64:

Query Match 66.6%; Score 1804.5; DB 2; Length 547;
Best Local Similarity 66.2%; Pred. No. 3.2e-77;
Matches 363; Conservative 75; Mismatches 107; Indels 3; Gaps 2;

OY 1 MASKELLDARKREKLSGVCKLANAVVYTGPKGRNVIEKSFSPVYTKGVSVAKEI 60
DB 1 MAKEIKFSTEARERKMLRGVDILANAVKATLCPKGRNVIDKSFAPRTTKGVSVAKEI 60
OY 61 ELEDKFENNAGOMVEKVAAPKTSIDAGDGTATVLAQAIYRGVVLVAGRPMAIKRGI 120
DB 61 ELEDKFENNAGOMVEKVAAPKTSIDAGDGTATVLAQAIYRGVVLVAGRPMAIKRGI 120
OY 121 DKAVAVAVTKELSDITRKPRDOKEIAOVGTISANSOTTTIGNITIAEMAKVKGCVITVEE 179
DB 121 DKAVAIAVAELKALSLVPSDFKAIAOVGTISANSKEICDIIAQAMEKGRNSGVITVEE 180
OY 180 AKGFTTLVVVEGMRKFDKGLSPFYTNPEKMCVCLDNPIILCEKTKISMKMLPILVEY 239
DB 180 GOSLENEEDLVVEGMRKFDKGLSPFYTNPEKMCVCLDNPIILCEKTKISMKMLPILVEY 240
OY 240 VAKVNPRLIIADVEGEALATLVVVKLRGALQOVAVKAPGCGERRKAMLEDIAITGCGE 299
DB 240 VAKVNPRLIIADVEGEALATLVVVKLRGALQOVAVKAPGCGERRKAMLEDIAITGCGE 299

RT "GroEL gene sequence of Vibrio parahaemolyticus."
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
CC -i- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CC CONDITIONS (BY SIMILARITY).
CC -i- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
CC 7 SUBUNITS (BY SIMILARITY).
CC -i- SIMILARITY: TO THE CHAPERONIN CPN60 (60KD SUBUNIT) DOMAIN.
DR EMBL: AF221845; AAF27528.1;
DR InterPro: IPR001844;
DR Pfam: PF00118; cpn60_Tcpl; 1.
DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
KM ATP-binding; Chaperone.
SQ SEQUENCE 548 AA; 57629 MW; E39223E289D6F7AC CRC64;

Query Match 64.7%; Score 1753; DB 2; Length 548;
Best Local Similarity 63.9%; Pred. No. 8.3e-75;
Matches 350; Conservative 87; Mismatches 109; Indels 2; Gaps 2;

OY 1 MASKEILFDKAREKLSRGVDKLANAVKVTGPKGRNVLEKSGSPVITKDGVSVAKEI 60
1 MAADVAFGNDARKMLEGVAVNLADAVKVTGLGKGRNVLDKSGADPTTKDGVSAVEI 60
DB 61 ELEDKFNMGQAQVYKVAAPRTSDIAGDGTATVLAQAIYREGVKLVAAGRNPAIKRGI 120
61 ELEDKFNMGQAQVYKVAAPRTSDIAGDGTATVLAQAIYREGVKLVAAGRNPAIKRGI 120
OY 121 DKAIVAVTKELSDITKTRROKELIAOVGTISANSDDTTIGNIIAEAMAKVGKGVITVEEA 180
121 DKAIVAVTKELSDITKTRROKELIAOVGTISANSDDTTIGNIIAEAMAKVGKGVITVEEA 180
DB 121 DKAIVAAVEQLKELSYECNDTKAIAQVGTISANSDSVGNIIAEAMERVRDGVITVEEG 180
OY 181 KGLFTTLDVVEGKFPDGRGVLSPYFVNPEKMCVLDNPYILNEKKTSMKMLPILEOV 240
181 KGLFTTLDVVEGKFPDGRGVLSPYFVNPEKMCVLDNPYILNEKKTSMKMLPILEOV 240
DB 181 QALQDELVDVEGQDFRGVLSPTFINQEGSVLEENPFLILDVKKISNIRELLPTLEAV 240
OY 241 AKVNRPLLIIADVEGEALATLVNKLRGALQVAAKAPFGERRKAMLEDIALLTGEA 300
241 AKVNRPLLIIADVEGEALATLVNKLRGALQVAAKAPFGERRKAMLEDIALLTGEA 300
DB 241 AKASRPLLIIADVEGEALATLVNMMRGIVKAAVAKAPFGDRKRAMLODIALITGCV 300
OY 301 IFEDRGIKLENVSLSLGAKRVVIDKENTTYDGAKESEDIAKAVKOIRAOIEETSSDY 360
301 IFEDRGIKLENVSLSLGAKRVVIDKENTTYDGAKESEDIAKAVKOIRAOIEETSSDY 360
DB 301 ISEIELELEKVTLEDLGAKKRVISITKENSTIIDGAGEEAMIOGRAVIRQOIEDATSDY 360
OY 361 DREKLOERLAKLVGAIVIHGAATEEMKEKKDRVEDALNTRAIVEGIVPGGTAFF 420
361 DREKLOERLAKLVGAIVIHGAATEEMKEKKDRVEDALNTRAIVEGIVPGGTAFF 420
DB 361 DREKLOERLAKLVGAIVIHGAATEEMKEKKDRVEDALNTRAIVEGIVPGGTAFF 420
OY 421 RSIKVLDIKPADDELGLAGNIIRSLLEPLRQIAANAGYEGSIYVEKREPKDGFENA 480
421 RSIKVLDIKPADDELGLAGNIIRSLLEPLRQIAANAGYEGSIYVEKREPKDGFENA 480
DB 421 RAASKIVDLEGEQNEQNGCIRVALRAMEAPIRQITKNAGDEDSVANNVAKAGEGSGTGYNA 480
OY 481 ASGEYEDLIKAGVIDPKKVTTRIALONAASVASILLTTCALAEKPEKKDMPGP-GGG 539
481 ASGEYEDLIKAGVIDPKKVTTRIALONAASVASILLTTCALAEKPEKKDMPGP-GGG 539
DB 481 ATGEYADLVGVNDPTKVTTRIALONAASVASILLTTCALAEKPEKKDMPGP-GGG 539
OY 540 GGMGMDGM 547
540 GGMGMDGM 547
DB 540 GGMGMDGM 547

RESULT 13
O9MWLA PRELIMINARY; PRT; 540 AA.

AC 09MWLA
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE 60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN).
GN GROEL.
OS Methylovorus sp. (strain SSI / DSM 11726).
OC Bacteria; Proteobacteria; beta subdivision; Methylophilus group;

OC Methylovorus
CX NCBI_Taxid=81683;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=SSI;
RA Eom C. Y., Kim Y. M.;
RT "Cloning, sequencing, and molecular analysis of the groEL operon from
RT Methylovorus sp. strain SSI DSM11726."
RL Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
CC -i- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CC CONDITIONS (BY SIMILARITY).
CC -i- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
CC 7 SUBUNITS (BY SIMILARITY).
CC -i- SIMILARITY: TO THE CHAPERONIN CPN60 (60KD SUBUNIT) DOMAIN.
DR EMBL: AF152236; AAD34149.1;
DR HSSP: P06139; IJON.
DR InterPro: IPR001844;
DR Pfam: PF00118; cpn60_Tcpl; 1.
DR PRINTS: PR00298; CHAPERONIN60.
DR PRINTS: PR00304; TCOMPLEXTCP1.
DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
KM ATP-binding; Chaperone.
SQ SEQUENCE 540 AA; 57399 MW; 0549FF7288713B62 CRC64;

Query Match 64.3%; Score 1741; DB 2; Length 540;
Best Local Similarity 66.6%; Pred. No. 3e-74;
Matches 359; Conservative 72; Mismatches 106; Indels 2; Gaps 2;

OY 1 MASKEILFDKAREKLSRGVDKLANAVKVTGPKGRNVLEKSGSPVITKDGVSVAKEI 60
1 MAKEVAFHDAHRTIRIVGVNLADAVKVTGLARAVMCLIERFSGAPVITKDGVSVAKEI 60
DB 61 ELEDKFNMGQAQVYKVAAPRTSDIAGDGTATVLAQAIYREGVKLVAAGRNPAIKRGI 120
61 ELEDKFNMGQAQVYKVAAPRTSDIAGDGTATVLAQAIYREGVKLVAAGRNPAIKRGI 120
DB 61 ELDCKFENMGQAQVYKVAAPRTSDIAGDGTATVLAQAIYREGVKLVAAGRNPAIKRGI 120
OY 121 DKAIVAVTKELSDITKTRROKELIAOVGTISANSDDTTIGNIIAEAMAKVGKGVITVEEA 180
121 DKAIVAVTKELSDITKTRROKELIAOVGTISANSDDTTIGNIIAEAMAKVGKGVITVEEA 180
DB 121 DKAIVAVTKELSDITKTRROKELIAOVGTISANSDDTTIGNIIAEAMAKVGKGVITVEEA 180
OY 181 KGLFTTLDVVEGKFPDGRGVLSPYFVNPEKMCVLDNPYILNEKKTSMKMLPILEOV 240
181 KGLFTTLDVVEGKFPDGRGVLSPYFVNPEKMCVLDNPYILNEKKTSMKMLPILEOV 240
DB 181 KSLDNELDVVEGQDFRGVLSPTFINQEGSVLEENPFLILDVKKISSIRDLPTLEAV 240
OY 241 AKVNRPLLIIADVEGEALATLVNKLRGALQVAAKAPFGERRKAMLEDIALLTGEA 300
241 AKVNRPLLIIADVEGEALATLVNKLRGALQVAAKAPFGERRKAMLEDIALLTGEA 300
DB 241 AKAGKPLLIIADVEGEALATLVNSMRGILKVAANKAPFGDRKRAMLEDIALLTGAIV 300
OY 301 IFEDRGIKLENVSLSLGAKRVVIDKENTTYDGAKESEDIAKAVKOIRAOIEETSSDY 360
301 IFEDRGIKLENVSLSLGAKRVVIDKENTTYDGAKESEDIAKAVKOIRAOIEETSSDY 360
DB 301 ISEIELELEKVTLEDLGAKKRVISITKENSTIIDGAGEEAMIOGRAVIRQOIEDATSDY 360
OY 361 DREKLOERLAKLVGAIVIHGAATEEMKEKKDRVEDALNTRAIVEGIVPGGTAFF 420
361 DREKLOERLAKLVGAIVIHGAATEEMKEKKDRVEDALNTRAIVEGIVPGGTAFF 420
DB 361 DREKLOERLAKLVGAIVIHGAATEEMKEKKDRVEDALNTRAIVEGIVPGGTAFF 420
OY 421 RSIKVLDIKPADDELGLAGNIIRSLLEPLRQIAANAGYEGSIYVEKREPKDGFENA 480
421 RSIKVLDIKPADDELGLAGNIIRSLLEPLRQIAANAGYEGSIYVEKREPKDGFENA 480
DB 421 RARSRIVNLKGGNGDODAGIRIVLRAIEAPALAIANAGDEPSVYINVLAGSGNFGYNA 480
OY 481 ASGEYEDLIKAGVIDPKKVTTRIALONAASVASILLTTCALAEKPEKKDMP-MPGG 537
481 ASGEYEDLIKAGVIDPKKVTTRIALONAASVASILLTTCALAEKPEKKDMP-MPGG 537
DB 481 ATGEYADLVGVNDPTKVTTRIALONAASVASILLTTCALAEKPEKKDMP-MPGG 539

RESULT 14
O33688 PRELIMINARY; PRT; 538 AA.

ID O33688
AC O33688
DT 01-JAN-1998 (Tremblrel. 05, Created)

DT 01-JAN-1998 (TRENBLREL. 05, Last sequence update)
 DT 01-MAR-2001 (TRENBLREL. 16, Last annotation update)
 DE 60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN) (FRAGMENT).
 GN GROEL.
 OS Sitophilus oryzae principal endosymbiont.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae.
 OX NCBI_TaxID=61800;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SER.
 RX MEDLINE-98042468; PubMed-9367844;
 RA Charles H., Heddi A., Guillard J., Nardon C., Nardon P.;
 RT "A molecular aspect of symbiotic interactions between the weevil
 Sitophilus oryzae and its endosymbiotic bacteria: over-expression of a
 chaperonin.";
 RT Blochem. Biophys. Res. Commun. 239:769-774(1997).
 ~- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
 PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
 CONDITIONS (BY SIMILARITY).
 CC -1 SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
 7 SUBUNITS (BY SIMILARITY).
 CC EMBL: AF0052236; AAB97670.1; .
 DR HSSP: P06139; 1JON.
 DR InterPro: IPR001844; .
 DR InterPro: IPR002423; .
 DR Pfam: PF00118; cpn60_TCP1; 1.
 DR PRINTS: PR00298; CHAPERONIN60.
 DR PRINTS: PR00304; TCOMPLEXTCP1.
 DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
 DR ATP-binding; Chaperone.
 FT NON-TRF 538 538
 SQ SEQUENCE 538 AA: 56683 MW: EFD17668C0F09D15 CRC64:

 Query Match 64.1%; Score 1736; DB 2; Length 538;
 Best Local Similarity 64.0%; Pred. No. 5, 1e-74;
 Matches 348; Conservative 81; Mismatches 109; Indels 6; Gaps 2;

DB 481 ATEEGNMIDKIDLPRTKVTSSALQYASISAGIMTTFECWYTD--QPEKDP----DLGC 534
 QY 541 MGM 544
 DB 535 MGM 538

 RESULT 15
 ID 09X6Y3 PRELIMINARY; PRT: 544 AA.
 AC 09X6Y3;
 DT 01-NOV-1998 (TRENBLREL. 12, Created)
 DT 01-NOV-1998 (TRENBLREL. 12, Last sequence update)
 DT 01-MAR-2001 (TRENBLREL. 16, Last annotation update)
 DE 60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN).
 GN GROEL.
 OS Bacteroides forsythus.
 OC Bacteria; CF8 group; Bacteroidaceae; Bacteroides.
 OX NCBI_TaxID=28112;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 43037;
 RA Reid H.I., Riggs M.P.;
 RT "Identification and nucleotide sequence of the heat shock protein 60
 (groEL) gene of Bacteroides forsythus.";
 RT Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.
 CC -1 FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
 PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
 CONDITIONS (BY SIMILARITY).
 CC -1 SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
 7 SUBUNITS (BY SIMILARITY).
 CC EMBL: AJ006516; CAB43992.1; .
 DR HSSP: P06139; 1GRL.
 DR InterPro: IPR001844; .
 DR InterPro: IPR002423; .
 DR Pfam: PF00118; cpn60_TCP1; 1.
 DR PRINTS: PR00298; CHAPERONIN60.
 DR PRINTS: PR00304; TCOMPLEXTCP1.
 DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
 DR ATP-binding; Chaperone; Heat shock.
 KW SEQUENCE 544 AA: 57971 MW: 4712157E0349494C CRC64:

 Query Match 63.6%; Score 1722; DB 2; Length 544;
 Best Local Similarity 63.5%; Pred. No. 2, 3e-73;
 Matches 345; Conservative 85; Mismatches 111; Indels 2; Gaps 2;

```

Db      362  KKKLLEERLAKLAGVAAYLYGAPSEVEKKEKKDVDAHLATRAIEEGVPGGAYLR  421
OY      422  SLKVLVDIKPADDELGLANTIRRSLEPTROIANAGYEGSTVEKVRPKOCFGNAA  481
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      422  AIPALEGKLGKENEDETTGIEIVKRAIEPTLQIYVNNNGKGAVVOKYEGTGAFGNAR  481
OY      482  SGEYEDLIKACVIVDPKRYTRIALDNASVASLLTTECATAEKREPKKMDPMGGGGM  541
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      482  TDVYEDLSEACVYDPAKVTRIALENASITAGMFLTTECVADKKEAPADPM-PCMGGM  540
OY      542  GGM  544
      :  :  :
Db      541  GGM  543

```

Search completed: August 1, 2001, 17:25:26
Job time: 432 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2001, 15:38:44 ; Search time 2206.1 seconds
(without alignments)

11547.701 Million cell updates/sec

Title: US-09-077-574A-1
Perfect score: 1647
Sequence: 1 atggcttctcaagaatacct.....gtatgacgtagtactag 1647

ring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_da1:*
- 2: gb_da2:*
- 3: gb_da3:*
- 4: gb_in1:*
- 5: gb_in2:*
- 6: gb_in3:*
- 7: gb_om:*
- 8: gb_ov:*
- 9: gb_pat1:*
- 10: gb_pat2:*
- 11: gb_ph:*
- 12: gb_pl1:*
- 13: gb_pl2:*
- 14: gb_pl3:*
- 15: gb_pl4:*
- 16: gb_da1:*
- 17: gb_da2:*
- 18: gb_fun:*
- 19: em_htgo_hum:*
- 20: em_htgo_inv:*
- 21: em_htgo_rod:*
- 22: em_htg_hum1:*
- 23: em_htg_hum2:*
- 24: em_htg_hum3:*
- 25: em_htg_hum4:*
- 26: em_htg_hum5:*
- 27: em_htg_hum6:*
- 28: em_htg_hum7:*
- 29: em_htg_hum8:*
- 30: em_htg_inv1:*
- 31: em_htg_inv2:*
- 32: em_htg_other:*
- 33: em_htg_rod:*
- 34: em_hum1:*
- 35: em_hum2:*
- 36: em_hum3:*
- 37: em_hum4:*
- 38: em_hum5:*
- 39: em_hum6:*
- 40: em_hum7:*
- 41: em_in:*
- 42: em_om:*
- 43: em_or:*

44: em_ov:*

45: em_pat:*

46: em_ph:*

47: em_pl:*

48: em_ro:*

49: em_sts:*

50: em_sy:*

51: em_un:*

52: em_v1:*

53: gb_sts1:*

54: gb_sts2:*

55: gb_sts3:*

56: gb_sy:*

57: gb_un:*

58: gb_v11:*

59: gb_v12:*

60: gb_v13:*

61: gb_v14:*

62: gb_v15:*

63: gb_v16:*

64: gb_v17:*

65: gb_v18:*

66: gb_v19:*

67: gb_v20:*

68: gb_v21:*

69: gb_v22:*

70: gb_v23:*

71: gb_v24:*

72: gb_v25:*

73: gb_v26:*

74: gb_v27:*

75: gb_v28:*

76: gb_v29:*

77: gb_v30:*

78: gb_v31:*

79: gb_v32:*

80: gb_v33:*

81: gb_v34:*

82: gb_v35:*

83: gb_v36:*

84: gb_v37:*

85: gb_v38:*

86: gb_v39:*

87: gb_v40:*

88: gb_v41:*

89: gb_v42:*

90: gb_v43:*

91: gb_v44:*

92: gb_v45:*

93: gb_v46:*

94: gb_v47:*

95: gb_v48:*

96: gb_v49:*

97: gb_v50:*

98: em_da3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1647	100.0	4710	3	LIU45241
2	798.8	48.5	2672	3	M86549 Amoeba prot
3	778.4	47.3	1811	3	M31918 Legionella
4	777	47.2	2715	1	AB057417 Pseudalt
5	771.2	46.8	1635	2	AB057417 Pseudalt
6	771.2	46.8	1894	2	M98257 Bartonella
7	770.6	46.8	1644	2	U78514 Bartonella
8	764.6	46.4	2241	3	AJ243594 Pseudalt

9	762.6	46.3	1644	3	ROU78515	U78515 Bartonella
10	756	45.9	2016	3	LPNHPB	M91673 Legionella
11	755	45.8	1627	1	AF014829	AF014829 Bartonella
12	751.2	45.6	2770	3	LMGROELSR	X57320 L. micdadei
13	744	45.2	2596	3	FTGRO	X58853 F. tularensi
14	741.8	45.0	96109	10	AX067460	AX067460 Sequence
15	738.6	44.8	1718	2	AF325222	AF325222 Streptococcus
16	732.4	44.5	5907	2	D85628	D85628 Buchnera
17	732.4	44.5	34699	1	AF008210	AF008210 Buchnera
18	725.8	44.1	11763	3	U32736	U32736 Haemophilus
19	722.6	43.9	2156	2	AF335323	AF335323 Listeria
20	721.2	43.8	2200	3	HEAGROESTL	M91030 Haemophilus
21	718.4	43.6	2468	2	CPHYRAB	X51404 Chlamydia
22	718.4	43.6	2718	2	CHTGRROESTL	L12004 Chlamydia
23	718	43.6	10890	1	AE006151	AE006151 Pasteurella
24	715.2	43.4	2577	2	CTU52049	U52049 Chlamydia
25	715.2	43.4	10461	1	AE002305	AE002305 Chlamydia
26	714.8	43.4	1647	2	AF221845	AF221845 Vibrio
27	714.8	43.4	2360	3	PMU30165	U30165 Pasteurella
28	714.4	43.4	1522	1	AF014833	AF014833 Bartonella
29	713.2	43.3	6224	1	AF003957	AF003957 Myzus
30	710.4	43.1	2275	2	CHTMRPAB	M1739 C. trachomatis
31	710	43.1	2739	2	APSYMSL	X6150 Acyrthosiph
32	710	43.1	347550	2	AP001118	AP001118 Buchnera
33	708.8	43.0	279110	3	RPXX03	AJ235272 Rickettsia
34	708.4	43.0	12560	1	AE006276	AE006276 Lactococcus
35	707.2	42.9	9993	1	AE001285	AE001285 Chlamydia
36	705.6	42.8	2621	3	RPY15783	Y15783 Rickettsia
37	705.2	42.8	1644	2	AY017169	AY017169 Vibrio
38	704.8	42.8	282183	2	CJ1168X4	AL139077 Campyloba
39	704.2	42.6	2309	2	BACSGROESTL	L10132 Bacillus
40	702.2	42.6	1980	2	AF321516	AF321516 Mycobacterium
41	701.4	42.6	296950	2	AF001508	AF001508 Bacillus
42	701	42.6	10389	1	AE006627	AE006627 Streptococcus
43	700.8	42.5	1647	3	CHTGRROESTL	M58027 Chlamydia
44	700.6	42.5	1647	3	RPY7380	U77380 Rhopalosiph
45	698.8	42.4	2267	3	LLGROESTL	X71132 L. lactis

ALIGNMENTS

RESULT 1
LOCUS L1045241 4710 bp DNA BCT 23-SEP-1998
DEFINITION Lawsonia intracellularis plasmid pISI-2 GTP-binding protein homolog
gene, partial cds, L27 50S ribosomal binding protein homolog, L21
50S ribosomal binding protein homolog, GroES/HSP10 homolog, and
GroEL/HSP60 homolog genes, complete cds.

ACCESSION U45241.1 GI:3599919
VERSION U45241.1
KEYWORDS

NAME Lawsonia intracellularis.
ORGANISM Lawsonia intracellularis
Bacteria; Proteobacteria; delta subdivision; Lawsonia.
REFERENCE 1 (bases 1 to 4710)
AUTHORS Dale, C.J.H., Moses, E.K., Ong, C.C., Morrow, C.J., Reed, M.B., Hasse, D.
and Strunell, R.A.

TITLE Identification and sequencing of the groE operon and flanking genes
of Lawsonia intracellularis: use in phylogeny
JOURNAL Microbiology 144 (Pt 8), 2073-2084 (1998)
MEDLINE 98386497
REFERENCE 2 (bases 1 to 4710)
AUTHORS Dale, C.J.H.

REMARKS Direct Submission
TITLE Submitted (11-JAN-1996) C. Jane H. Dale, Molecular Biology,
Victorian Institute of Animal Science, 475 Mickleham Road, Attwood,
VIC 3049, Australia
FEATURES Location/Qualifiers

SOURCE 1. 4710
/organism="Lawsonia intracellularis"
/plasmid="pISI-2"
/db_xref="taxon:29546"

CDS	complement(<1..345)	
	/codon_start=1	
	/transl_table=11	
	/product="GTP-binding protein homolog"	
	/protein_id="AAC36496.1"	
	/db_xref="GI:3599922"	
	/translation="MRFVDEVTITSVAGKGGNCVAFRRKFKPKGPNCGDGGN IIFKADSRLLTLDFVQRRIYRANQEGGSGRHKGGEDLLHLPVGTITFEOLD KEFLVLDPRGY"	
	complement(508..774)	
	/codon_start=1	
	/transl_table=11	
	/product="L27 50S ribosomal binding protein homolog"	
	/protein_id="AAC36497.1"	
	/db_xref="GI:3599921"	
	/translation="MAHKAGCGSSRRGDSRPGRCITKRFGGKVLAGNIVRLQGT IHPGENVGMGRDYLTPAKVAGIYETVETFRKCRYSRVHIQPTV"	
	complement(787..1095)	
	/codon_start=1	
	/transl_table=11	
	/product="L21 50S ribosomal binding protein homolog"	
	/protein_id="AAC36498.1"	
	/db_xref="GI:3599922"	
	/translation="MYAIEAGCKQFCVEGSKIFVSKIDAEGTEIFDKIFMGGS SPQIGTPYINNAKVIKLEHGRDKILVEKKRRNDSRLQGRDYLAKVTGTLQ "	
	1595..1900	
	/note="heat shock protein; chaperonin"	
	/codon_start=1	
	/transl_table=11	
	/product="GroES/HSP10 homolog"	
	/protein_id="AAC36499.1"	
	/db_xref="GI:3599923"	
	/translation="MNLKPLNDRLVLRLESEKTAGLYIPTAKKESRGRVAVG PGKHDDGKLLPMAVAKGDTVLEKNYAGTEVKKIDGVEHLVREDIYLAIVTGTGRK 2048..3694	
	/note="heat shock protein; chaperonin"	
	/codon_start=1	
	/transl_table=11	
	/product="GroEL/HSP60 homolog"	
	/protein_id="AAC36500.1"	
	/db_xref="GI:3599924"	
	/translation="MASKELFDPAKREKLSRGVDLANAVKVTIGPKGRNVIEKSF GSPVTRKDSVAKKEIEIDKPEFNAGQVAKVAPTSIAGGTTATVYLAQIIRE GKVLNAGRNPAIKRGIDRAVAATVKEISDITKPRDKEIAQVGTISANSPTIGN IIEAKAVKGGVITVEAKGLLETLVDVEGKFPDGLSPFTVNPPEKMCGLDNP YILCNKPKITSMKMDLPILEOVAKVRPLIIAEDVEGALATLVNKLGAQVAV KAPGFERKAMLEDIAITLGEAIFEDRGIKLENSLSLGAHVVDIKEDNTIVD GAKSESDIRARVQIRAOIETSSDREKLOELAKLVGAVVHVGATETSEK KDRVEDALNATRAVVEGIVPGGCTAFVRSIVLDDIKPADDELAGLNIIRSEK LRQIANAGTEGSIYVEKREPKDQCFNNAASEYEDLIKAGVIDPKKVTRIALONAA SVASLLITTECAIAEKPEPKDMPGCGGCGMDGM"	
	3821..4702	
	/note="orf; similar to Escherichia coli possible protoporphyrinogen oxidase (hemK)"	
	/codon_start=1	
	/transl_table=11	
	/protein_id="AAC36501.1"	
	/db_xref="GI:3599925"	
	/translation="MIRYIDMIGKTTMFRKAGIDSPKLSAEILISVINTRLQIIM TPPEPIPTSSYLNDIMLRGEPYIAVLTGKKEFSRFKQTQATLLPRETELLI EPLNHNIPQOYIFADIGTGGSCGIIITAAEKKNMLGINTDSSALATKNSKN NHSQLOFLOSDFQPLCPSSLDLITSNPYSISENLTSLEPVIISFEPKALTPAK CIHLDEINVLHCYKRIITQAEISLPGGIIILEHQAQAEAILLLKNNIWTNVISH TDLTKNNRFTYAKYKI"	

BASE COUNT 1638 a 827 c 785 g 1460 t
ORIGIN
Query Match 100.0%; Score 1647; DB 3; Length 4710;
Best Local Similarity 100.0%; Pred. No. 5; 3e-276;
Matches 1647; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 atgagctctaaagaaatcccllllgaagctaaagcccgtagaaacattccagagtgta 60
 Db 2048 atgaccttcaaaagaaatccctttttatgcttaaaccccgcaaaaactttcacagctgta 2107
 Oy 61 gataaactctgcaaaagcgttgtaagtaaaccttggaacctaaagccgtaagtgcttatt 120
 Db 2108 gataaacttgcaaaagcgttgtaagtaaaccttggaacctaaagccgtaagtgcttatt 2167
 Oy 121 gaaaagctcttggtctcccgatgtatcaaaagatgtgtatctcgttgcaaaagaaatt 180
 Db 2168 caaaagctttgttggtctcccgatgtatcaaaagatgtgtatctcgttgcaaaagaaatt 2227
 Oy 181 gaacttgaagaataagtttgaagaataatggcgctcaaatggttaaaagagtagctcccaa 240
 Db 2228 gaacttgaagaataagtttgaagaataatggcgctcaaatggttaaaagagtagctcccaa 2287
 Oy 241 actagagatattcgtgtgtgtggaactcaacagcaacagtccttgcaagctatttat 300
 Db 2288 actagagatattcgtgtgtgtggaactcaacagcaacagtccttgcaagctatttat 2347
 Oy 301 cgtgaagctgttaaaactgttagcagctgtgtatctcctatgagcattaaagtgagata 360
 Db 2348 cgtgaagctgttaaaactgttagcagctgtgtatctcctatgagcattaaagtgagata 2407
 Oy 361 gataaagctgt 420
 Db 2408 gataaagctgt 2467
 Oy 421 caaaagaaatagctcaagctgtggaacatctctgcaaacctgtatcaacaatagtaac 480
 Db 2468 caaaagaaatagctcaagctgtggaacatctctgcaaacctgtatcaacaatagtaac 2527
 Oy 481 atcagagctgaagctgaagctgaagctgaagctgaagctgaagctgaagctgaagctga 540
 Db 2528 atcagagctgaagctgaagctgaagctgaagctgaagctgaagctgaagctgaagctga 2587
 Oy 541 aaaagctctgaaactataatgaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 600
 Db 2588 aaaagctctgaaactataatgaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2647
 Oy 601 tctcgaactctgtgaactcaaacctgtggaagaaatgtgtgtgtgtgtgtgtgtgtgtgtgt 660
 Db 2648 tctcgaactctgtgaactcaaacctgtggaagaaatgtgtgtgtgtgtgtgtgtgtgtgtgt 2707
 Oy 661 ctctgaatggaagaaatgactgacgaagaaatgactgacgaagaaatgactgacgaagaaat 720
 Db 2708 ctctgaatggaagaaatgactgacgaagaaatgactgacgaagaaatgactgacgaagaaat 2767
 Oy 721 gctaaagctaaacgctcactcctatctatctgctgaagaaatgactgacgaagaaatgact 780
 Db 2768 gctaaagctaaacgctcactcctatctatctgctgaagaaatgactgacgaagaaatgact 2827
 Oy 781 acaactgtgactaataaagctccgctggaagcaactcaagctgtgagccgtaaaagctccgct 840
 Db 2828 acaactgtgactaataaagctccgctggaagcaactcaagctgtgagccgtaaaagctccgct 2887
 Oy 841 ttggtgtaacgctgaagaaatgactgacgaagaaatgactgacgaagaaatgactgacgaagaa 900
 Db 2888 ttggtgtaacgctgaagaaatgactgacgaagaaatgactgacgaagaaatgactgacgaagaa 2947
 Oy 901 atatttgaagaaatgactgacgaagaaatgactgacgaagaaatgactgacgaagaaatgact 960
 Db 2948 atatttgaagaaatgactgacgaagaaatgactgacgaagaaatgactgacgaagaaatgact 3007
 Oy 961 aaagctgtgactaataaagctccgctggaagcaactcaagctgtgagccgtaaaagctccgct 1020
 Db 3008 aaagctgtgactaataaagctccgctggaagcaactcaagctgtgagccgtaaaagctccgct 3067
 Oy 1021 gataatgaagctgaagctgaagctgaagctgaagctgaagctgaagctgaagctgaagctga 1080
 Db 3068 gataatgaagctgaagctgaagctgaagctgaagctgaagctgaagctgaagctgaagctga 3127
 Oy 1081 gataatgaagctgaagctgaagctgaagctgaagctgaagctgaagctgaagctgaagctga 1140

Db 3128 gatcctgaaaaaacttcaagaacgcttctcaaaaactgttgctgagctgattatccat 3187
 Oy 1141 gttgagctgtctactgaacacgaaatgaaagaaagaaatgaaagaaatgaaagaaatgaaagaaat 1200
 Db 3188 gttgagctgtctactgaacacgaaatgaaagaaagaaatgaaagaaatgaaagaaatgaaagaaat 3247
 Oy 1201 aatgcaacaagagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1260
 Db 3248 aatgcaacaagagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 3307
 Oy 1261 cgtccatataagctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1320
 Db 3308 cgtccatataagctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 3367
 Oy 1321 aatgcaacaagagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1380
 Db 3368 aatgcaacaagagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 3427
 Oy 1381 gaagctctatgt 1440
 Db 3428 gaagctctatgt 3487
 Oy 1441 gcatcaggaagatataagaaactataaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1500
 Db 3488 gcatcaggaagatataagaaactataaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3547
 Oy 1501 cgtatgcatataaagaaatgcaagcaatgcaagcaatgcaagcaatgcaagcaatgcaagcaatgca 1560
 Db 3548 cgtatgcatataaagaaatgcaagcaatgcaagcaatgcaagcaatgcaagcaatgcaagcaatgca 3607
 Oy 1561 attgctgaaaaaactgaaacacaaatgaaacacaaatgaaacacaaatgaaacacaaatgaaacacaa 1620
 Db 3608 attgctgaaaaaactgaaacacaaatgaaacacaaatgaaacacaaatgaaacacaaatgaaacacaa 3667
 Oy 1621 atgggtgataagacgataagctag 1647
 Db 3668 atgggtgataagacgataagctag 3694
 RESULT 2
 XBCGROEX
 LOCUS XBCGROEX 2672 bp DNA BCT 02-APR-1998
 DEFINITION Amoeba proteus symbiotic bacterium groes and groEL genes, complete cds.
 ACCESSION M86549
 VERSION M86549.1 GI:2996603
 KEYWORDS
 ORGANISM
 SOURCE Amoeba proteus symbiotic bacterium.
 ORGANISM Amoeba proteus symbiotic bacterium.
 Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group; Legionellaceae.
 REFERENCE 1 (bases 1 to 2672)
 AUTHORS Ahn,T.I., Leeu,H.K., Kwak,I.H. and Jeon,K.W.
 TITLE Nucleotide sequence and temperature-dependent expression of XgroEL gene isolated from symbiotic bacteria of Amoeba proteus
 JOURNAL Endocyt. Cell Res. 8, 33-44 (1991)
 REFERENCE 2 (sites)
 AUTHORS Ahn,T.I., Lim,S.T., Leeu,H.K., Lee,J.E. and Jeon,K.W.
 TITLE A novel strong promoter of the groEX operon of symbiotic bacteria in Amoeba proteus
 JOURNAL Gene 148 (1), 43-49 (1994)
 MEDLINE 95011657
 REFERENCE 3 (bases 1 to 2672)
 AUTHORS Jeon,K.
 TITLE Direct Submission
 JOURNAL Submitted (05-FEB-1992) Kwang W. Jeon, Zoology, University of Tennessee, Cumberland Street, Knoxville, TN 37996-0810
 COMMENT On Mar 30, 1998 this sequence version replaced gi:155400.
 FEATURES
 source
 1..2672
 /organism="Amoeba proteus symbiotic bacterium"
 /db_xref="taxon:2728"


```

Oy 788 tagtcaataagctccgtagagacacccaagtgttagccgttaaaagcttcgtgttg 847
Db 818 TAGTCAACACACTGCGCGGTATTGTAAGAATGCTGTCAAGCGCTGGTTTGGG 877
Oy 848 aagcgcgttaagatctgtctgaagatctgtatcccttaagtggaagaagaatattg 907
Db 878 ATCGTCGCAAGCGATGTTGCAAGACATTGCTATTGCTAAGAGGTCAAGTATTTCG 937
Oy 908 aagatcgtgtataaagcttgaanaatgtgaagctgtctctttaggaacagctaaag 967
Db 938 AAGAAATTTGGCAAGAGCGCTTGAAGGTGCTACTCTGGAAGATCTTGTAGTCCAGCA 997
Oy 968 tagtattgacaanaaataactactatcgtttagatgctgagcgcggaataacagaatata 1027
Db 998 TCGTTGTTACCAAGAAAGAAACACTATATCATTTGATGTAAGGAAGCAACTGAAATTA 1057
Oy 1028 aagctcgaatttaacaactctgtgcacaacttgaagaacaagctcagatataatgctg 1087
Db 1058 ATGCTCGTATTGCTCAATTTGCTGCACAAATGGAAGAAACCACTTCTGATTACGATTAG 1117
Oy 1088 aaaaacttgaagaacgtcttgcanaaacttctgtgtgagtagctgtatccatgttgag 1147
Db 1118 AAAAATTTACAAGACGCGCTTCTTAAGTTAGCTGCTGCTGCTTCTGTTATCAAAAGTTGCG 1177
Oy 1148 ctgtctactgaaacttgaatgaagaagaagaagatcgtgttagaagatctcctaataatgcaa 1207
Db 1178 CTGCCACACAAAGTTGAATGAAGAAGCAAGAAAGCGTTCGTTGAATATCTCTTCATGCTA 1237
Oy 1208 caagcgcgtcgttgaagaagaagatctgtccctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1267
Db 1238 CTCGCGCTGCAAGTGAAGAAAGTATCGTTCGCGGTGGTGGCTTCTGATTGCTGCTC 1297
Oy 1268 ttaaaagcttgaatataaaccctgtgaatgaatgaatgaatgaatgaatgaatgaatgaat 1327
Db 1298 AGAAAGCTCTGATTCATTGCAAGCGATGAATGACGATCAAAATATCGGTATCATATT 1357
Oy 1328 tccgtcgtctcttgaagaagcctttagcctaactgtgcgaatgtgtgtgtgtgtgtgtgtgt 1387
Db 1358 TAGCTCGCGCTATTGAAATCTCAATGCGTCAAGATTGTTACTACCAAGATATGAAGCTT 1417
Oy 1388 ctattgttgaagaagaatctgtgaacccaagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1447
Db 1418 CTGTTGTGTAACAGAGTACTAGACCAAGAACAACTAGGTTTCAACGCTGCAACTG 1477
Oy 1448 gagaatcgaagaaccttaagaagctgtgtcatgtatgaccttaaaagttacaagttatgt 1507
Db 1478 GTCAATACCGTATGCTGATGAGATGGCTATTCTTGATCAACAACTAACCCGTTANCG 1537
Oy 1508 catlaaanaaagcagcatcagtagctctctactctactactactactactactactactactact 1567
Db 1538 CTCTGCAAAATGACGCTTCTTAGCGCACTTGTGATCTTACTACTGTAATGCTTGTCTG 1597
Oy 1568 aaaaaccagaacctaaagaatagatgctatgcctcgtgcgtgtgtgtgtgtgtgtgtgtgtgt 1627
Db 1598 A-----TCTGCTTAAGAAAGAAAGAGGTGTTGTGCGCGGATATGGCGGATATGGCGG 1651
Oy 1628 gtatgagcgtcatg 1641
Db 1652 GAATGGCTGCATG 1665

```

```

RESULT 4
AB057417 2715 bp DNA BCT 16-MAR-2001
LOCUS AB057417 Pseudalteromonas sp. PSIM3 gene for GroES, GroEL, complete cds.
DEFINITION AB057417
ACCESSION AB057417
VERSION AB057417.1 GI:1336171
KEYWORDS
SOURCE Pseudocalteromonas sp. PSIM3 (strain:PSIM3) DNA.
ORGANISM Pseudocalteromonas sp. PSIM3
Bacteria; Proteobacteria; gamma subdivision; Alteromonadaceae;
Pseudocalteromonas.
REFERENCE 1 (sites)

```

```

AUTHORS Kurusu, Y. and Nakamura, T.
TITLE Molecular chaperone of Psychrophile
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2715)
AUTHORS Kurusu, Y.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-2001) Yasuou Kurusu, Ibaraki University School
of Agriculture, Laboratory of Molecular Microbiology, Chuo 3-21-1,
Ibaraki Aml, Ibaraki 300-0393, Japan
(E-mail: krusu@ipc.ibaraki.ac.jp, Tel:81-298-88-8646,
Fax:81-298-88-8646)
FEATURES
source
1..2715
location/Qualifiers
/organism="Pseudalteromonas sp. PSIM3"
/strain="PSIM3"
/db_xref="taxon:87791"
/country="Japan"
/note="Isolated from the Japan Trench"
promoter
-35_signal
376..388
promoter
402..410
-10_signal
406..411
RBS
458..464
/gene="groES"
458..761
gene
474..761
/gene="groES"
474..761
CDS
/gene="groES"
/codon_start=1
/transl_table=1
/product="GroES"
/protein_id="BAB39464.1"
/db_xref="GI:1336172"
/translation="MNIRPLQDRVIVKRELEETKASGIVLTGSAAKSTRGEVAVG
NGRIIDNGDVRALEVKAGDVFVFCSEYKTEKTEKIQGYLIMREDNIIIGV"
797..803
gene
797..803
RBS
811..2460
CDS
/gene="groEL"
811..2460
/gene="groEL"
/function="molecular chaperone"
/codon_start=1
/transl_table=1
/protein_id="BAB39465.1"
/product="GroEL"
/db_xref="GI:1336173"
/translation="MAKEVLFAGDARAKMLTGVNIIANVKTIGPKGNVYIDKSF
GSPITTKDQSVAKETLEDEKFEKMGQKQKQEVASKANDAGDGTTRATVLAOSTIVE
GLKAVAGNMPMDLKRIGIDKAIINAVALKALSVPCADTKAIAOVGTISANSKEIGD
ITAEAMKVRNMGCVIVVEGOSLENDLVVEGMPDRGYLSPFTINNAKGAVEIDN
PFILLVKKSINIRELLPTLEAVAKASKPLIIAEDLEGALATLVNNKRGVIVSA
VKAPGFGDRRKAMLODIATVLTGTVISEIGLELEKATVGLAKKVVITKDTFTI
DGAGEEGGICGVSOIKQIJEATSDVDKDEKLBRKATGAGVAVIKVATLEKME
KKDRVEBALHATRAVVEEGYVPGGVALVYRAKSLVELLDNEDONHGVAVLABRA
PLRQIVTNAGDEASVIVNAVKGEGNGVNAATEYINDMIEMLDTPTKVTRSLQFA
GSIAGLMTTEAVVAELIPKEBAACAPMGCMGCMGCMW"
2519..2567
terminator
BASE COUNT 860 a 513 c 613 g 729 t
ORIGIN
Query Match 47.2%; Score 777; DB 1; Length 2715;
Best Local Similarity 67.4%; Pred. No. 2.9e-15;
Matches 1111; Conservative 0; Mismatches 535; Indels 3; Gaps 1;
Oy 1 atgctcttaagaatactcttttgatgtctaaagccgtgaaaacttccagaggtgta 60
Db 811 ATGGCAGCTAAAGAGTACTTTTTCACGTACGACGCGCTGTAAGTCTTACTGCGCTA 870
Oy 61 gataaacttgaaatcgtttaaagtaacacttgaacctaagagccgtgaatcgtctat 120
Db 871 AACATCCTAGCAACGCGATTAAAGTTACATTAGTCTCTAAAGCCGCTAACGTTGTACTA 930

```



```
/db_xref="GI:397596"
/translaction="MAAKPVKPCRDAAERLLRVDILADAVKVTLPGRKRVYIDMSF
GAPRTIKDGVSAKLELEKLEFENMGQOMREVASKSTNDIADGCTTATVLGAIVDE
GVRAVAASMPDLKRGIDAEVAADVLKAKKIDTSEIIOVAVIISANGAEDIOK
MIADAMEKVNIEGIVTEVEKTAETEELEVEGQDFEAGVPEVTESEKMDLDP
YILIEHEKSLNLOSILPVEVAOSGKPLIIEADVEGALATLVYKLGKIAV
KAPGFDRRKAMLEIDIAVLTSGOVISEDVICTLENTYEMLAGAKRVHSEKETTIVD
CAGOKSEINRVSQIKQIETTSDDYRELOERLAKAGVAVIRVGSSTEVYEVK
KDRVDDALNATRAAVEGIVPAGGTPLLRAKALUSTKGNRPDEAGIGILRLAQARA
ROIANAGAEAAVIVCKVLENCSDTFEGYNATVADPRLISFQIVDPKVVRSALONAA
SIASLTTEEMAYEVPKEAAMPAMGCGMGDMF"

BASE COUNT      471 a      283 c      425 g      456 t
ORIGIN

Query Match      46.8% Score 771.2; DB 2; Length 1635;
Best Local Similarity 68.3%; Pred. No. 3.1e-14;
Matches 115; Conservative 0; Mismatches 508; Indels 9; Gaps 3;

OY 1 atgagcttaagaatccttcttgaagctaaagcccgtaagaacttcacgaagtgta 60
DB 1 ATGGCTGCTAAAGAACTAAATTTGGCCGTCACGACCTGAAAGCTGTTFACGTCGTG 60
OY 61 gataaacttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttga 120
DB 61 GATATCCTTGCTGATGCTGATTAAGGTAACTTGGTCTTAAGGCCGCAATGCTGAT 120
OY 121 gaaaagcttcttgcctccagcttatacaaaagatgtagtctgttgcacaaagaa 180
DB 121 GATAAATCATTTGGTGGTCCACGCACTTAAAGATGCTGTCGCTGCAAAAGAAATC 180
OY 121 gaaacttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttga 240
DB 181 GAACTTGAAGAACTGTTGAATATGCGCCGCAAAATGTCGCTGAAGCTTCCTCAAA 240
OY 241 actcgaatcttgcctgaagcttgaagcttgaagcttgaagcttgaagcttga 300
DB 241 ACCAAGCATTTGCTGCTGATGCAACGACACCAACTGTTTAAAGCCAGCATCTT 300
OY 301 cgtgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagct 360
DB 301 CAGGAAGCTGTAAGAGCTTGGACGCGATGTAATCAATCAATCTTAAAGCGGAAT 360
OY 361 gataaagcttgcctgaagcttgaagcttgaagcttgaagcttgaagcttga 420
DB 361 GATGCTGCTGTTGAAGCGGTTGTCACATCTTTTCAAAAAGCGAAAGATCAAAAC 420
OY 421 caaaagaaatgaagcttgaagcttgaagcttgaagcttgaagcttgaagct 480
DB 421 TCAGAGAAATTCACAAAGTTCATTTCTGCTAATGGCCCTGAGACATCGCTAAA 480
OY 481 atcagctgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttga 540
DB 481 ATGATGCTGATGCTGTAAGAAAGTCCGATGATGATGATGATGATGATGATG 540
OY 541 aaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttga 600
DB 541 AAACAGCTGAACCTGACTGTAAGCTTGTGAAGAAATGCTGATGCTGCTGATCT 600
OY 601 tctcaatcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagct 660
DB 601 TCTCCTTATCTTGTTCACCAATTTCTGAATAAATGATGCTGATGATGATGAT 660
OY 661 ctctgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagct 720
DB 661 CTATATCTGTAAGAAATTTGCTAATCTTGTGCTGCTGCTGCTGCTGCTGCT 720
OY 721 gctaaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagct 780
DB 721 GCTCAGCTGCTGAACCTCTTCTCATTAATGCTGAGATGATGAGAGGCTTTTG 780
OY 781 acaacttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttga 840
```

```
DB 781 ACAGCTGCTGTAACCAAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
OY 841 ttctggaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagct 900
DB 841 TTTGGTGAACCGCTGAAGCTATGCTGGAAGATATGCAAGTTTGACTTCTGCTG 900
OY 901 atattgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagct 960
DB 901 ATTCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
OY 961 aaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagct 1020
DB 961 AAAAAGTGCATGCTTCTTAAGAAACCAACATGCTGCTGCTGCTGCTGCTGCT 1020
OY 1021 gataaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttga 1080
DB 1021 GAATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
OY 1081 gctcgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttga 1140
DB 1081 GACCTGAGAAATGCAAGAAAGCTTCAAAACCTTGTGCTGCTGCTGCTGCTGCT 1140
OY 1141 gctcgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttga 1200
DB 1141 GTTGTGCTTCAACAGAAAGTGAAGTGAAGAAAGAAAGAAAGAAAGAAAGAA 1200
OY 1201 aatgcaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttga 1260
DB 1201 AACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
OY 1261 cgtcgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttga 1320
DB 1261 GCTGCTGCAAGAGCTCT--TTCTATTAAGAAAGAAAGAAAGAAAGAAAGAAAG 1320
OY 1321 aatcgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttga 1380
DB 1318 GGTATCATTCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
OY 1381 gaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagct 1440
DB 1378 GAACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
OY 1438 gctcgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttga 1497
DB 1438 ACACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1497
OY 1498 aacgcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagct 1557
DB 1498 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1557
OY 1558 gctcgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttga 1614
DB 1558 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1614
OY 1615 gctcgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttga 1670
DB 1618 GCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1670

RESULT 6
BAOBB63A      1894 bp      DNA      BCT      26-Apr-1993
LOCUS      Bartonella bacilliformis immunoreactive protein (bb63) gene,
DEFINITION      complete cds.
ACCESSION      M98257.1 GI:143845
VERSION      M98257.1
KEYWORDS      Immunoreactive protein.
SOURCE      Bartonella bacilliformis (library: ATCC 35685) DNA.
ORGANISM      Bartonella bacilliformis.
REFERENCE      1 (bases 1 to 1894)
```

AUTHORS Xu, Y., Lu, Z. and Thier, G.
TITLE Molecular cloning, expression and sequencing of extracellular Bb63,
JOURNAL a immuno-reactive protein released by Bartonella bacilliformis
Unpublished (1992)
FEATURES
Location/Qualifiers
1..1894
/organism="Bartonella bacilliformis"
/db_xref="taxon:774"
/tisue_1lb="ATCC 35685"
RBS 58..62
/gene="Bb58"
58..1755
/gene="Bb58"
70..1704
CDS
/gene="Bb58"
/codon_start=1
/transl_table=11
/product="immunoreactive protein"
/protein_id="AA22898.1"
/db_xref="GI:143846"
/translation="MAKEVFGDREBRLRGVDILDAVKVTLGPKRNVYIDSKF
GAPRTKGVSAKEIELENKFNENGAOMLEEVASKNDIAGDGTATVGOAVOE
GKVAASAMNMDLKRIDAENAVNADLFKKAKIOTSEELAOYATISANAEIDGK
MIADMEKVNQEVITVEAKTAETELVBSMOPDRGLSPFYVNSEKMMVDLDP
YILHEKRLNSQLPLVLEAVOSGKPLLIADVEGELATLVNKLRGGLTAAY
KAPGGRKRAMLEDIAVITSGOVISERVGILLENVLEMLGRKRVHSEKTTIVD
GAGKSEINARVSOIKOIEETSDYDREKLEOLAKLGVAVRVGSEVEYERK
KDRVDDANATRAVAEGLVPGGTALLRAKALSIKKNPDQEGIIIRALDAPA
ROIANNGEAAVIVGKLENSDPTFGNTATACGDLISGIVDPVAVNSALONAA
SIASLLITTEAAVEVPKKEAAPAMPGGMGGMF"
terminator
1727..1755
/gene="Bb58"
1748..1788
stem-loop 557 a 326 c 476 g 535 t
BASE COUNT
ORIGIN

Query Match 46.8%; Score 771.2; DB 2; Length 1894;
Best Local Similarity 68.3%; Pred. No. 3e-124;
Matches 1115; Conservative 0; Mismatches 508; Indels 9; Gaps 3;

QY 1 atgagctcctaaagaatccctcttgatgctaaagccgctgaagaacttcacgaagtgta 60
DB 70 ATGCGCTGCTAAAGTAATTTGCCCTGACGCGCAAGGTTGTTACGTCGTGTC 129
61 gataaactgcaaaagctgtaaaagaacttgacacaaagccgaagctgctgatt 120
DB 130 GATATCTGCTGATGCTCTTAAGTAACACTTGTCTTAAGGCGCGCAATGTGCTATT 189
QY 121 gaaaagctcttgctcccaagtattacaagaagtgtgctatctgtgcaaaagaat 180
DB 190 GATTAATCATTTGGTCTCCACGCAATFACAAAAGATGTTGTCCTGCAAAAAGAAATC 249
QY 181 gaacttgaagaagtctgaaataatgtggcgctcaaaagtgttaagaagtgcctcaaa 240
DB 250 GAACCTTGAACACAGTTTAAAAVAVGGCCGCCCAATCTTGCCTCAAGTGGCTCAAAA 309
QY 241 actagagatattgcctgtagaagaacttacaagaacagctcccttcagaagctattat 300
DB 310 ACCAACGATATTGCTGCTGATGACACACACAGCAACTGTTTTAGGCGACGCGCATCTT 369
QY 301 cgtgaagctgaagaactctgacagctgtgcttaactcctatggccttaagaagctgata 360
DB 370 CAGGAAGGTGTGAAGCTCTGTGACCGCAGTATGATCAATGATTTAAAGCGCGGAATT 429
QY 361 gataaagctgtgtgtgctgcttactaaagaactaagcgacattacaagcctactgtgac 420
DB 430 GATGCTGCTGTGGAACGCTGTGTTGTCGATCTTTTCAAAAAACGGAAGATCAACAAC 489
QY 421 caaaaagaagaacttcaaaatgaaaccttctgcaaaactctgatacaacaatagataat 480
DB 490 TCAGAGAAATTCACACACTTATTTCTGCTAATGGCGCTGAGGACATCGGTAAA 549

QY 481 atcatagctgaagctatggtctaaagttggaagaagtgctatcaacagcttgagaagct 540
DB 550 ATGATGCTGATGCTATGCAAAAAAGCGGTATGAGAGGTATACACGCTAGAGAACGA 609
QY 541 aaaagcttgaagaactatagatgtgtgtgaagaagaatgaagttgacggtgctaccc 600
DB 610 AAAACAGCTGAACCTGATGAGTGTGTTGAAGGAATCAGTTGATGTGGGTATCTT 669
QY 601 tctccactctgttaactatccctgagaagaatggtctgtgtaacttgataccctatc 660
DB 670 TCTCCTTACTTTGTCCACCAATTCGAAAAAATGATGTTGATCTTATCATCTTACATC 729
QY 661 ctgtgaatgagaagaagaatctactgcatgaagaagacatgctacccaacttgagaaga 720
DB 730 CTTATTCATGAAGAAATTTGCTAATCTGCAATCTTGTGCTGCTGCTGAGCTGTT 789
QY 721 gctaaagtaaacgctccactccttattatgtctggaagaagtgagaagtgagaacttga 780
DB 790 GTTCAGTCTGTAAAGCTCTTCTCATATTCCTGGAAGATGTGAAGGTGAACCTTGGCA 849
QY 781 aaactgtatgcaataaagctccgtggaagcaccctcaagctgtgacgctgaagaagctcgt 840
DB 850 ACCTGCTGTGTAAACAAAGTTGGTGGTTTAAATTCCTGCTTAAAGCTCCAGCT 909
QY 841 ttgtgtaagcgcgtgaagaactatgcttgaagaatgtctatccctactggaagaaga 900
DB 910 TTGGTGACCGCGGTAAACCTATGTGGAAGATATGACGTTTGTACTTGTGTCAGGTG 969
QY 901 atatttgaagaatcgtgtttaaagcttgaagaatgtgaactgtctctcttgaagaagct 960
DB 970 ATTTCTGAAGATGTGGTATTAAAGCTTGAAGAAATCTTAACTTAACTTAACTTAACT 1029
QY 961 aaacgtgtatgtatgacaaagaagaataactactatcgtgtgtaagtgctggaagaatcaga 1020
DB 1030 AAAAAAGTCAATGTTCTTAAGAAACCAACCAATGTTGATGCTGTGACAAAAGCTT 1089
QY 1021 gatattaagctcgaagttaacaagaatcgtgacaaacttgaagaagaagaactcaat 1080
DB 1090 GAATCAATGACAGCTGTATGACAGATTAGGACAGATTGAAGAAACCACTTGTGACTAT 1149
QY 1081 gatcgtgaagaacttcaagaagaagctcttgaagaactgtgtgtggaatgactgtatc 1140
DB 1150 GACCGTGAGAAATGTGAAAGAAACACTTGAAGAACTTGTGCTGCTGCTGCTGCTGCT 1209
QY 1141 gttggaagctgctactgaagaactgaagaagaagaagaagatcgtgtgaagaatgctca 1200
DB 1210 GTTGGTGTTCACAGAAAGTTGAAGTGAAGAAAGAAAGACCGCTTATGACCGCTTG 1269
QY 1201 aatgcaacaagaagctgctggttgaagaaggtatgtccctgtgtgtgtgtgtgtgt 1260
DB 1270 AACGCAACAGCTGCTCTGTAGAGAAAGGATGTTCTGCTGCTGCTGCTGCTGCTGCT 1329
QY 1261 cgtccatcaagaagctctgtatgatatataccgctcgatgagatgaactgtgtgaact 1320
DB 1330 CGGTGCAAGAGCTCTTCTTATTAAGGAAAGCAACCAATCAATCAAGAGGATATT 1386
QY 1321 aatatcctcgtcgtctctcttgaagaagccttcaagcaatgtcgaagaatgtgtgtat 1380
DB 1387 GGTATCATCCGCGCTGCTTGCAGAGCTCCAGCAAGCTGATGCTATATGAGGTGA 1446
QY 1381 gaaggtctcatgtgtgtgaagaagaagctcgtgaa---ccaagaatgtgttgaatla 1437
DB 1447 GAAGCAGCGGTATTGTTGGCAAGAGTGTGAAGAAATTTGTCAGATCATTTGGTTCAAT 1506
QY 1438 gctgcatcaggagaataatagaagccttataaagctgtgtgtcatctgataccataaag 1497
DB 1507 ACAGCAGCGCTCAATTCGTTGATTTGATTTCTTTGGATTCCTTATACAGTGAAG 1566
QY 1498 acacgtatgcattcaaaaatgagcatcagatgagctccctactcttaactacagagatgc 1557
DB 1567 GTGCGTGTGCTGTGAAATGACGATCAATGCAAGTCTTTGATGACAGCAAGAACG 1626
QY 1558 gctattgtcgaaaaaaccagaacctaaagaataatgacct--atgctcgtggtgtgtatg 1614

[illegible]

Db	121	GATTAATCATTTGGTGGCCCTCCCATTTACAAAGATGGTATCCGTTGCAAAAGAAATTC	180
Oy	181	gaacttgaagaataagcttgaaaaatattgggcgcacaaatbgttlaaaagaagtaagctcccaaa	240
Db	181	GAACTGGGAAGTAAGTTCCGAATAATATGGGTGCCCAATATGCGGTGAACCTTCTCTPAA	240
Oy	241	actcgcgatalctgcttgttgaatgaactacaacacgaacgcgcctctgcacaagctattat	300
Db	241	ACCATGACATTTGGTGGTATGGACAAACACGGCAACTGTTTATGAGACAGGCTATTGTC	300
Oy	301	cgtaagagtgtaaaabactctgtagaagctggttcgttaatccctatggtccataaagctgata	360
Db	301	CAAAAGAGGTGTAAGAAAGCCGTTTGGTCGAGGATGAAACCCAAATGATCTCAAACTGGCTATC	360
Oy	361	gataaagctgtgtgtgtgttctactaaagaactaagcgacattacaagaagcctactcgtgac	420
Db	361	GATGCTGCTGTTGATGTAAGTGGCGCAAAATCTTTTCAAAAAGCCGCAAAAATTCCAACT	420
Oy	421	caaaagaagaataagtcgaagcttgaacacattctgcacaactcgtatacaacaataagtaat	480
Db	421	TCACAGAAATATGCAAGATGGGACAAATTTCCGGTAATAGTGGTGTGTAATCGGCAAA	480
Oy	481	atcatagctgaagctatgctctaaagtttgaaaaaggagtgtatcaacagtttgaaggaagt	540
Db	481	ATGATCCCTGATGTCATGCAAGAAAAGTTGGCAATGAAGCGTGTATTACCGTTGACAAACCA	540
Oy	541	aaaggtctgaanaatcatatagaatgtgtttgaagaagaataagttgaacgttgctactc	600
Db	541	AAAACTGCTGAAGCAATTAAGAAAGTCGTTGAAGGAAATCACTTTGATGTGGATATCTT	600
Oy	601	ctccatacttctgtaactaaatccctgagaagaatgttttctggaacttgaataccctatac	660
Db	601	TCGCCATTACTTGTGCACCAATGCTGAAGAAATGGTGGCTGTGATGATGATCCTTAACTT	660
Oy	661	cttctgaatgagaaaaagattactatgcatagaaagacatgctacccaactctagaacaagt	720
Db	661	CTTATTTCACGAAAAGAAACCTCTATATTCGCCAATCCCTTCCTTCACGATCTTGAACCTCTT	720
Oy	721	gctaaagtaaacgcctcaaccccttatattgttcgaagaacgtgaagaagtgaaagcacttga	780
Db	721	GTTTACGTCGTGTAAACCATCTTCTCATTTATTCCTCGAAGATGTGGAAGGTGAACCTTGGCA	780
Oy	781	aaactctgtagtcaataaagctcgttggagacatcccaagcttgcgcgtlaaaagcctcctg	840
Db	781	ACGCTCGATTGCAACAAACTGCGTGGGTTGMAAATTGTGCTGTGAAGGCTCCAGGA	840
Oy	841	tttgggtgaagcgcgtlaaagctatgcttgaagatatgtataccctactcgtgaaggagaaga	900
Db	841	TTTTGGTATGCGCCCTAAAGCCATGTAGAAAGATATTGCATATTAACTCCGCAACGCTT	900
Oy	901	atatttgaagaatcgttgtataaagcttgaanaatgaagcttgcctctttaggaacagct	960
Db	901	ATTTCCTGAAGTGTGGCATTTAAATTTGAAAATATGTCACTTTGGATATGCTTGTGCTGTGA	960
Oy	961	aaacgtgtagttatctgcaagaagaataactactatcctgttgcattgctcgtgaaataacagaa	1020
Db	961	AAGAAAGTCATATTTCTTAAGAAGAAATATCACCATATTATGATGATGGTGGCAAAAAGAGC	1020
Oy	1021	gatatlaaagctcgtatgtaaacaatcgtbgaacaatctgaagaacaagctcagatattat	1080
Db	1021	GAAATTTAATCGCCCGCTCAACCAATTCAGAGTACAGATGTGAAGAAAACAACTTCCTACATAT	1080
Oy	1081	gattcgttaaaaacttgaagaagcgtcttgcaaaactgtgtgtgtgagtgatgcgtatcat	1140
Db	1081	GACCGTGAATAATATGCAAGAAAGACTTGTCTAACTCTGCTGGAGGTCTTCTGTTATTCCT	1140
Oy	1141	gttgaagctcgtactcgtgaacaactgaatgaagaagaagaagatcgttgaagaagatgctcta	1200
Db	1141	GTTGGTGGAGCAACGAAGATTGAAGTGAAGAAAGAAAGATCGTGTGATGATGATCCCTTG	1200
Oy	1201	aatgcacaagaagagctgcgttgaagaagaagbatgtccctcgtgtgtgtactgtcttgtc	1260

OY	241	atcagagatattgctgvgatgagaaactatacaagcaacgcgctctctgacaagaactatctat	300
Db	744	gctaacgatacagccggcggaatgatacaactaccctacactactctgacagctattgta	803
OY	301	cgatgaaggctataaactctgtagcagccggtctgtaacctctgcccataaactgtagcata	360
Db	804	aacgaagccgtttaaagcgtttgacagcagcattgaccccaattgacacctttaaagccggcattc	863
OY	361	galaaagctgltgtgtctgtactataaagaactaagcgacatliaaagcctactcgtgac	420
Db	864	gacaaagcagtaaatgtgctgctgttgacagactttaaagcgttatctgttcacatgctctgac	923
OY	421	caaaaagaatagctccaaagttggaaccaatttcgcgaaactctgatacaacaatagtgtaac	480
Db	924	actaaagaatttgacacagatgagatcattttacgctaaactctataaagatgttgccgac	983
OY	481	atcaatagctggaagctatgagcctaaagcttgcgaaagagagtgltataccagttgagaa	537
Db	984	atcattttgcacacagcaatggaanaagtagccctgacagtggtgtttatctgtagaaaga	1043
OY	538	gctaaaggctctgaaactacataatagatgtgttgaaaggaatgaaagctttagccgtgctac	597
Db	1044	gctcaatcactagaaatgaaactgcgattgttgaaagcattgacagtttaccgtgtgattac	1103
OY	598	ctctctccataactcttgaactaatccctgaaanaagctgtgtgaaactgtataacccttat	657
Db	1104	ctattctcttcttcttattataacagcccgaaaaaaagctagcttgaaactatgataccctatt	1163
OY	658	atccctgtgataagaaagaatgatacagatagaaagaacatgctcccaatcttataaaca	717
Db	1164	attctactttagatataaaaaaatatgtctaacatttgcgcaattttttaccaacattttaaaca	1223
OY	718	gttgcataaagtaaacccgttccactccatattatgtcgaagaagctgaaagtgaaagcactt	777
Db	1224	gttgctaaagcaaacagcagccactgataatgtgcgaagaaagacctgaaagcgaacattta	1283
OY	778	gcaacactgtgtagctcaataagctccgtgtgagcaactccaaagctgtgagccgaaagctctc	837
Db	1284	gctcattattagatgataataacatgcctggtattgttaaaagatgtacagctgaaacactcct	1343
OY	838	gattctgtggaagccgttaagactatgcttgaagatattgctatccctactcagaaagaa	897
Db	1344	ggttttgccacgcctgttaaaccaatgtctaacagatgtcgtgttttaactgtgcggtact	1403
OY	898	gcaaatcttgaagatcgtgtatataaagcttgaagaatgtaagctgtctctctttaaagaca	957
Db	1404	gtgatttcagaaagaatgcgcttggacttgaaagcttgaanaagcaacagttgaagacctgagta	1463
OY	958	gctaaacgtgtgattatctgcaagaagaataactactatcgttgaatgtgtctgaaataca	1017
Db	1464	gctagagcgcttaattatcacataaagatgtatcacacaaatcatgtgatgtgctgccaacaa	1523
OY	1018	gaagaatctaaagctcgaagttlaaacaaatctgtgcacaaatgtgaagaacaagaactcagat	1077
Db	1524	gcacgatttttaacggccctgttttcacaaattttaaagccaatgtgaagaagcaacacacagac	1583
OY	1078	tatagatcgtlaaaaaactcgaagaacgtctctgaaaaactgtgtgtggaagtgcgttata	1137
Db	1584	tacgattaaagaaaaaacacacagacagccatgtgctaaactagctggcggtgttcagtaattc	1643
OY	1138	cagtgtgagagctgtactcgaagaactgaatgaaagaaagaaagatcgtgtgaagaatgtct	1197
Db	1644	aaatttttgctgctgcactgaaatgaaatgaaagaaagaaagaaagcggctgtgaagatgca	1703
OY	1198	ctaaagtcacaagaagctcgggttgaagaaggtattgtccctcgtgtgtgtgaactgctttc	1257
Db	1704	tttaatttcacactctgcagcggttgaaagaggtgtatgtaacctgcggcgcggtgttcacta	1763
OY	1258	gtccgcctcacaatlaaagctccttgaatgatactaaacctgtgatacgtgaactgtcgtga	1317
Db	1764	gttgcgtcagcaagatgtaagctggtgtattttagttggcgacacagaaagattcaaaacacagct	1823
OY	1318	cttaataatccatccgtcgtctccttgaagaagcctttaaagcaaatgtctgcaaatgtctgc	1377

	DB	1824	ATTAAGTTGCACCTTGTGTCGAATGAAGCGCCACTTGTCGAATATCTAACCAAGCAGCT	1883
	OY	1378	tatgaaggtcttacttttgttagaaaaagtctgtgaacccaaagaatggttttgatttaac	1437
	DB	1884	GACGAAGCTTCACTTCTAATTAAATGACGAGTTAAAGCGGTTGACGTAACCTTTGGTAAACAAT	1943
	OY	1438	gctcacaicaggagaataitaagacccttatcaaacgcgtgcgtcatcatcctaataaaaagt	1497
	DB	1944	GCTGCTACTGCGGAATACATCATGATGATGGAATGGGTATCTCTAGATCTCAACTAAAGTA	2003
	OY	1498	acaagctatcgattacaacaatgcagcataagtagccttcacttcctaactcaacagaatgc	1557
	DB	2004	ACGCCGCTCTACCAATAATTTCCAGCGCTTATTGGCGCGGTCTAATGATCACCTACCGCAACT	2063
	OY	1558	gctatcgctgcgaataaacccagaaccttaaaaaagatatgcccattgccgcggcggtgatggt	1617
	DB	2064	ATGCTGACTCGAAA-----TCCCTAAAGATGATACGCCGCTCTGATATGGTGCGCATGGCG	2117
	OY	1618	ggatagtggtgcatgacgagcgtatgcatca 1646	
	DB	2118	GGTATGGGTGGAATGGCGCGGTATGATGTA 2146	
RESULT	9			
LOCUS	R0U78515	1644 bp	DNA	BCT 06-SEP-1997
DEFINITION				Bartonella quintana heat shock protein HSP60 (groEL) gene, complete cds.
ACCESSION	U78515			
VERSION	U78515.1	GI:2358235		
KEYWORDS				
SOURCE				
ORGANISM				Bartonella quintana. Bartonella quintana Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; Bartonellaceae; Bartonella. 1 (bases 1 to 1644) Haake,D.A., Summers,T.A., McCoy,A.M. and Schwartzman,W. Heat shock response and groEL sequence of Bartonella henselae and Bartonella quintana Microbiology 143 (Pt B), 2807-2815 (1997)
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
SOURCE				
gene				
CDS				
BASE COUNT				

ORIGIN

Query Match	46.38;	Score 762.6;	DB 3;	Length 1644;
Best local Similarity	67.88;	Pred. NO. 9.5e-123;		
Matches 113;	Conservative	0;	Mismatches 519;	Indels 9; Gaps 3

[illegible][illegible]

```

gene      /db_xref="taxon:446"
          /gene="hnpb"
          /gene="hnpb"
          /codon_start=1
          /transl_table=11
          /product="58-kDa common antigen"
          /protein_id="AA02598.1"
          /db_xref="GI:149690"
          /translation="MAKELRFQDDARLQMLAGVNALADAVOTMGPGRGKTVLEKSYK
APVYKQGVSAKEIEIEFHPFRPMNQAOVKKVSAKSTPADGVTYTAIVLARSILVEGE
HAAEAERKQKGVITVEDGNLBNELSVSEMOILAIVSPFIINQOMSCLEHPF
LVLVDKRSIEMSLVEGVAKSGRPLIADVEDEGAHSLPFIINQOMSCLEHPF
PGFDEKRAMLDQIDAILKRGVIVISEIKSLDELDGSAKRIYVKEMTTIDG
GKAIEINARIQIIRAOMETTSDYDREKQSRVAKLGVAVIKVGAATEMEKRR
RVEADLHTRAAVEGIIAAGGVALIRQKALDSIKGNDONNGINTLREAIISPM
QIVINACYEASVYVNVKVAEHKDNQGFNAARCEYGDPMVEKGLDLPKTYTRALQNAASY
ASLMLTTCMHADLPKKEEGVAGADMGGMGMGMGMGM"

BASE COUNT      591 a      323 c      508 g      594 t
ORIGIN
jerry Match      45 98; Score 756; DB 3; Length 2016;
est Local Similarity 67.3%; Pred. No. 1,3e-121;
Matches 1100; Conservative 0; Mismatches 525; Indels 9; Gaps 2;

Qy      8 ctaaagaacatccttttctatgctaaagcccgctgaaaaacttccagaggtgataaac 67
Db      215 CTAAGAATTTACGTTTGTGTCATGCCTCGCTCAAAATGCTTGCCTGATTATGCAAT 274
Qy      68 ttgcaaatgctgtttaaagtaaacacttggaccttgaagccgtaaatgtctgtatgaaagt 127
Db      275 TAGCAGATGCGGTTCAAGTACTATGAGGCCACGTGCGTATATGTTGGAATAAT 334
Qy      128 ctttggctcccgagctattacaagaagtgtgtctctgttgcataaagaacttgaactg 187
Db      335 CTTAGGGCGCTCTACTACTTAACTAAGACGGGTGCTCTGTTGCCAAATAATTTGAGTTTG 394
Qy      188 aagataagtttggaaatgtggcgcttcaaatgtgttaagaagtaagtaagctccaaactaagc 247
Db      395 AOCATCGTTTTCATGACATGGGCGCTCAAAATGGTTAAAGAAAGTGGCTTCAAAACTTCG 454
Qy      248 aatattcgtgtgaltgaaactcaacaagaagaagtcctctgcacaagctattatcgtgaag 307
Db      455 ATACGTGCGGTGATGTGTACTACTACTGTGACACAGTATTGGCTGTCTATTCTGTGTAAG 514
Qy      308 gttgaaactctgttagcagctgtgctgtaactcctaagtcacatlaaacgtgtgcataagaa 367
Db      515 GTCACAAAGCGAGTTGCTGCTGTTGATGATCCAAATGATGATCTCAAAACGGGTATTGTAAG 574
Qy      368 ctgtgtgtgcgtgttactaagaagaactaaagcgagcatlaacaagcgctactctgtgacaaag 427
Db      575 CAGTATTAGCGAGTTTACCAAAATTTACAAAGCTATGCTTAACCCATGCAAAACACCAAG 634
Qy      428 aaatgctcaagtttggaaaccttctcgtgaaactctgtatacaacaagaataatcatcag 487
Db      635 CTATGCGTCAAGTTGGAGACTATTTCTGCTAATTCGATGAGAACGATTGGTCTATCATTTG 694
Qy      488 ctgaaagctatggtctaaagtctggaaaagagatgttatcacagtttgaaggaagctaaagctc 547
Db      695 CTGAAGCAATAGAAAMAATGTTGGTTAAAGGGGTGTTATTAACCGTGAAGATGGTAAATGAT 754
Qy      548 ttgaaacacacttgaagtggtgttgaagaagatgaagtttgaacggtggcgtaactctctcat 607
Db      755 TGGAAATGACCTTCTCTGTTGTAAGGTATGCAATTTGATC--GCGGTAACTTCTCCAT 811
Qy      608 accttgaacacatccctgagaanaatglttctgaaactgtataaacaccttatactccgttga 667
Db      812 ACTTTATCAACACACGCAAAACATGAGACTGTGAACCTTGACATCCATCTCTTTATTGG 871
Qy      668 atggagaaaagataactcagcatgaaagacatgctaccaaactcttgaacaagttgctaaag 727

```

LOCUS	AF014829	1627 bp	DNA	BCF	20-MAY-1999
DEFINITION	Bartonella henselae 60 kDa heat shock protein (groEL) gene, partial cds.				
Db	872	TTGACAAAAAGTTTCACGATATTCGTGAATCTTGTCCGATTTGGAAAGCTTGCCAAAT	931		
Qy	728	taaacgccaccctcccttatattgctggaagacgtaagaagctgaaacactgcaaacctg	787		
Db	932	CTGGTCGTCTCTTTATTGATCATTTGCGAAGATATTTGAAAGCGCAACTTTAGCTACCTTGG	991		
Qy	788	tagtcaataagctcgttggaagcactcacaagtgtgaagcgaagaagctccggttttggt	847		
Db	992	TAGTCACAACAATGCGCGGATTTGTAAAGTATGTGCTGTCAAGAGCCCTTGTTTGCTG	1051		
Qy	848	aaagcgttaagctatgcttgaagatctgctacccctctgtaggaagaacaatattg	907		
Db	1052	ATCGCGGCAAGCGATGTTCGAAGACATTCCTATTTTGACTAAGGCGTCAAGTTATTCTCG	1111		
Qy	908	aagatctgtggttaaaagcttgaagaatgtaagctgtcttcttlaagaaacgctaaacgtg	967		
Db	1112	AAGAAATTTGGCAGACCTTTGGAAGGTGCTACTCTTGAAGATCTTGGTATGTCAACCGAA	1171		
Qy	968	tagtattacacaagaanaatactactacatcgctgtagtctgctggaatacagaagatca	1027		
Db	1172	TGCTGTGTTCCAAAGAAACACACTACATCATTTATGTGTGAAGAAAGCACTGAATTA	1231		
Qy	1028	aagctcgaattcaacaatctgctgcaaatggaagaacaagctcagattatgactcgtg	1087		
Db	1232	ATGCTCGTATTTACTCAAAATTCGTGCACAATGGAAGAAACCACTTGATTACGATAGAG	1291		
Qy	1088	aaaaattcaagaagaagctctgcaaaactgtgtgtagaagtaagctgtatccatgttgaag	1147		
Db	1292	AAAAATTTACAAAGCGCGCTTGCTAACTTGCTGCTGCTTGTCTATCAAAAGTTGGCG	1351		
Qy	1148	ctgctactgaaacctgaatggaagaagaagaagatcgctgaagaagtgctctaaatggaa	1207		
Db	1352	CTGCTACAGAAATTTGAAATGAAGAAAGAAAGACACGTTTGAGATGCTCTTCATGCTA	1411		
Qy	1208	caagagctgcggttggaagaagctatgtccctggtgtgtgtaactgcttctgtccgcctca	1267		
Db	1412	CTCGCGCTGCATAGAAAGAGGTATGCTTGCCGCTGTGCTGCTGCTGCTGCTGCTGCTG	1471		
Qy	1268	ttaagctccttgatgatlatataaactgtcgtatgatatgaataactgtcgtgacttaalca	1327		
Db	1472	AGAAAGCTCTTGGATTCATTGGAAGGCGGATTAATGACGATCAAAATATGCTTCAAAATTT	1531		
Qy	1338	tcgcgcgtctctctgtaagagcctcttaagcgaactgtcgaagtgctgaagctgaaggtt	1387		
Db	1532	TACGTCGGCGCTTTTAAATCTCCAATGCGTCAATTTGTTACTTAACGAGATATGAAGCTT	1591		
Qy	1388	ctattgtgtgaanaaagcttcgtgaaaccaaaagaatgatttgatttaagctgacacag	1447		
Db	1592	CTGTTGTAGTAAACAAGAGTACCTGAGCAAAAGACACACTACGCTTCAACGCTCCACACG	1651		
Qy	1448	gagaatatagaagaaccttaaaagctgltgcatlgatgactcaaaaaaglttaacagtattg	1507		
Db	1652	GTAATACGGCTATTTGCTGTGAGTGGGATATCTTGATCTGATCAACCAAAATTAACCCGATAG	1711		
Qy	1508	cattcaaaatgcaacatcagtagagctctcctactccttaactcagaagaatgcgtattgctg	1567		
Db	1712	CTCTCAAAATGACACTTCTGTAGCCAGTTTATGTTGACTGACTGAATCTATGTTGCTG	1771		
Qy	1568	aaaaacagaacctaaanaagatactgactatgccctggtggtgtgtagtctgtagtggtg	1627		
Db	1772	A-----TCTGCGCTTAAGAAAGAAAGAGGTGTGGTGGCGCGCATATGGCGGCATATGGCG	1825		
Qy	1628	gtatggaagcgtatg 1641			
Db	1826	GAATGGTGGCATG 1839			

ACCESSION	AF014829
VERSION	AF014829.1
KEYWORDS	GI:3603158
SOURCE	
ORGANISM	Bartonella henselae. Bartonella henselae Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
REFERENCE	Bartonellaceae; Bartonella.
AUTHORS	1 (bases 1 to 1627)
TITLE	Martson,E.L., Sumner,J.W. and Regnery,R.L.
JOURNAL	Evaluation of interspecies genetic variation within the 60 kDa
MEDLINE	heat-shock protein gene (groEL) of Bartonella species
REFERENCE	Int. J. Syst. Bacteriol. 49 Pt 3, 1015-1023 (1999)
AUTHORS	2 (bases 1 to 1627)
TITLE	Martson,E.L.
JOURNAL	Direct Submission
REFERENCE	Submitted (21-JUL-1997) Viral and Rickettsial Zoonoses Branch,
FEATURES	Centers for Disease Control and Prevention, 1600 Clifton Road,
SOURCE	Mailstop G13, Atlanta, GA 30333, USA
Location/Qualifiers	
gene	1..1627
CDS	/organism="Bartonella henselae" /strain="Houston-1" /db_xref="taxon:38323" 111..>1627 /gene="groEL" 111..>1627 /gene="groEL" /codon_start=1 /transl_table=11 /product="60 kDa heat shock protein" /protein_id="A0D04238.1" /db_xref="GI:3603159" /translation="NVIDSGFSGAPRIRKDGVSAAKEIEDEKFENMGAMIREVASKT NDADGGTTATYVGQIVQGVKAVAGAMPDLKRIDAAYDEVYANLEFKRAKTIO TSNEAOVGTISANGAEIIGKMIADEAKNGNCVITVEPAKTAETVAEVEGNOPR GYLSPFVTNAEKKNVADLDPIYLHEKKLSINOSLLPYLEAVVQSKPLLI1AEVDK GEALATI.VVNRKLKGCLIAAVKAPGFEDRRKANMEDIALITSQVISEDVIGKLENTV LDMGRKKVNISKRENTIIDVAGOKSEINARVOIKVQIEFTSDYDRKLOERLAK LAGGVAIVRVGATAEVEKEKDVRDALNATRAAVEGIIVAGGTALLLRANALTVK GSPNDAGGINIVRBALQAPAROAIATNAGEEMAIIVYKVLNNADPTGYMTAGEFD LIAIGTVPDKVVRSALONMASIASLLITTEMAVYAEVKPDTPVPPRMGGMCGMC"
BASE COUNT	482 a 306 c 413 g 426 t
ORIGIN	
Query Match	45.8% Score 755; DB 1; Length 1627;
Best Local Similarity	68.1%; Pred. No. 2e-121;
Matches 1111; Conservative	0; Mismatches 510; Indels 10; Gaps 4;
OY	1 atggcctctaaagaatcccttcttgatgactaaagcccgtagaaaaactttcaagagtgta 60
Dd	1 ATGGCTCTAAAGAAGTCGAAGTTTGCCCGTAGAACGCCGCTGAGCCGTTTTCGCCGCCTT 60
OY	61 gataaactgcgaatgcgtgttaaagtacaccttcygaacctaaaggccgtaattcgattat 120
Dd	61 GATATCCTTCTCAACGCGTGTAAAGTAAACACTCGGCCCTTAAGG-CGCANTGTGATGC 119
OY	121 gaagaagctcttggttccccagctatcataaaaagaatgtagtatcgttgtcaaaagaaatc 180
Dd	120 GATTAATCATTTTGGTGGCCCCCGCATTTACAAGAAGATGTTGATCCGTTGCCAAAAGAAATTC 179
OY	181 gaacttgaagaataagcttcttgaaatatagggccgtccaacttcygttaaagaatgagctcccaa 240
Dd	180 GAACCTGAAGATTAAGTTTCGAATAATATATGGGTGCGCAAAAGTGTGCGGAAGTGTCTTCTTAA 239
OY	241 actgagcatatcgtctgtgatctgaactacaacagcaaacagcttccttgcaaacagattat 300
Dd	240 ACCAATGACATTGCTGTGATGATGAACAACAACGCACTGTTTAGGACAGCATATTGTC 299
OY	301 cgtagaagctgtaaaaaacttgtagcagctcgtgctgtaattcctaaggccataaacgttggcata 360
Dd	300 CAAGAAGCTGTAAAGCCGTTGCTGTCACAGCCATGAACCACCAATGGAATCTCTCAAACGTCGATTC 359

QY	361	gataaagctcttcttgctgttactaaagacaaagacattcaaaagctactctgtgac	420
Db	360	gATCCTCTGTTGTGTAAGAGGTGGCAAAATCTTTTAAAAAAGCGAAAAATTCAAACT	419
QY	421	caaaagaataagctcaagcttggaaccattctgcacaactcgtatacaaaatagtaat	480
Db	420	TCACACAAATTCGACAAAGTGGAACAAATTTCCGCTAATGCGTCTGTGTAATTCGGCATA	479
QY	481	atcatacgtgaagcctatgagcttgaagaaagagtgatatacaagcttgaagaaagct	540
Db	480	ATGATTCCTGATGCCAATGAAAAAGTTGGCAATGAAGGTGTGATTACCTTGAAGAACA	539
QY	541	aaaggtcttgaactaatatagatgtgtgttgaagaaatgaagcttggacggtgctcaactc	600
Db	540	AAAACCTCTTAACCGAATTAATGAAAGTCGTTGAAGAAATCAAGTTGATGTGATATCTT	599
QY	601	ctcccaactcttctaactaaaccttgagaagaaatggtcttgaactctgtaaccctatctc	660
Db	600	TTCCCTTACTTTGTTCACCAATGCTGTGAATAATGTGTGCTGATCTTGATGATCTTACAT	659
QY	661	cttctgtaatagaaaaagatctactgaacatgaaagacatgcttcaaccaacttgaacaagt	720
Db	660	CTTATTCAGAAAAAGAAACTCTGTATATCTGCATCTCTGCTTCCAGTACTTGAACGCTT	719
QY	721	gctaaagtaaacccgtccactcctctatattgtctgaagacgtgaagatgtaagcacttgc	780
Db	720	GTTTCAGCTGTGAACCACTCTTCATATCTGCTGAAGATGTGAAGGTGAAGCTTTGGCA	779
QY	781	aaactgttagtaataagcccgcttgagaacccaagttgtgacgctgaagaaagctctcgt	840
Db	780	ACGCTCTTTCACCAACTGCGTGGTGGTTTGAATATGCTGTGTGAACCTTCACGA	839
QY	841	tttggtaaacgacgtaaaagctatgccttggaatatgtcatctcttacccttaccggaagaaagca	900
Db	840	TTTGGTATCGCGGTAAAGCAATGTGAAGAATATGTGAATCTTAAATCCGGACAGTT	899
QY	901	atatcttgaagatcgtgtgtataaaagcttgaanaatgttaagcttgccttcttgaagaagct	960
Db	900	ATTCTTCMAATATTGGCAATTAATTGGAATAATGTCACTTTGGAATATGTTGGTGTGCA	959
QY	961	aaacggttagttcttgcaaaagaaatactactatcgttgaatgagtgctcggaataacga	1020
Db	960	AAGCAACTCAATATTCTTAAGAAATACCGATTAATGATGAGTGTGGACAAAAAGAC	1019
QY	1021	gatactaaagctcgtatlaaacaactctgtgcacaatltgaagaacaagaactcagatatt	1080
Db	1020	GAATTAATTCGGCGGCGTCAACCAATCAAGATGATGCAATGGAAGAAACAATTTGATCAT	1079
QY	1081	gatcgtgaaanaactcaagaagctctgtcaaaaactgttgtgtgaagtagctgtatccat	1140
Db	1080	GACCGTGAATAAATTTGCAAGAAACACTTGTCTAACTGCTGGAGGTGTGCTTATTCGT	1139
QY	1141	gttggagactgtctctgaactgaatgaagaagaagaagatcgtgtgaagaatgactcta	1200
Db	1140	GTTGGTGGACCAACAGAGTTGAAGTGAAGAAAGAAAGAAATGATGATGATATCCCTTG	1199
QY	1201	aatgacaagaagctcgtgttgaagaagattgtccctcgtgtgtgtgaactgttctg	1260
Db	1200	AAGCAACAGTGTGACACTGTGGAAGAAAGATATTGTCGGCGTGTGGAACCCGATTTG	1259
QY	1261	cgtctcaataaagctccttgatgatactaaacctgtcatgatatgaactgtcgtgact	1320
Db	1260	CGTGCACAAATCTCCCTGACTG---TTAAAGGAAGCAATCTCGATCAGGAAGCTGTATC	1318
QY	1321	aataatacgtgtgtctctcttgaagagcctttagtcaaatgtcgaagaatgcttggcat	1380
Db	1317	AATTAATGTTGCTGTGCATCTCCAAAGGCCAGACAGCAATTCATCAAAATTCGCGTCAA	1376
QY	1381	gaaggtctctatctgtlaaanaagcttgtagccaaa---agaatgttttgaattaat	1437
Db	1377	GAAAGACGATTTATCGTTGGCAAAAGTCTTTGAAGAAACAATGCAGATACATTTTCCTTACAC	1436

```

Oy      1438 gclgcatcgaaggaatatgaagaaccttataaagcttgtygtatattgatccctaataaagt 1497
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1437 ACCGCACCCGGGGAATTTGGTGATTGTTCCTTGGGAAATGGTGCATCTGTGAAGCTT 1496
Oy      1498 aaacgatctgcatacaaatcgagcatcagtcagtcgcccttacttaacttaactcaagaatgc 1557
        || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1497 GTGCCCTCCTCGCTTCAACAATTCCTCCCTCAATTTGCTAGCCTTCTCATTTAACACACAGA 1556
Oy      1558 gctatctgtgaanaaacccaagaccataaaaagatalgcct--atgcctggcggltgtagt 1614
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1557 ATGCTTGGCTGAAGTTCCCAAGAAACACACCAGTGCCTCCAAATGCCTGTGGCGAATG11616
Oy      1615 ggtagatcagg 1625
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1617 GGTGGAAATGGG 1627

RESULT 12
LMGRELSR          LMGRELSR       2770 bp   DNA           BCT         18-FEB-1992
LOCUS              L.micdadei DNA for gro ELS operon .
DEFINITION X57520
ACCESSION X57520
VERSION X57520.1 GI:44099
WORDS       groELS operon.
            Tatlockia micdadei.
            Tatlockia micdadei
            Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
            Legionellaceae; Tatlockia.
            1 (bases 1 to 2770)
REFERENCE
AUTHORS Hindersson,P., Holby,N. and Bangsborg,J.
TITLE    Sequence analysis of the Legionella micdadei groELS operon
JOURNAL  FEMS Microbiol. Lett. 77, 31-38 (1991)
FEATURES
source    1..2770
            /organism="Tatlockia micdadei"
            /db_xref="taxon:451"
BASE COUNT 846 a 520 c 675 g 729 t
ORIGIN
Query Match 45.6% Score 751.2; DB 3; Length 2770;
Best Local Similarity 67.6%; Pred. NO. 8.4e-121;
Matches 1104; Conservative 0; Mismatches 518; Indels 12; Gaps 3.
Oy      8 cttaaagaatcccttttgtatgctlaaagcccgctgaanaaaccttcacgaggtgataaac 67
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      472 CTAAGAATTAAGTTTGGCGACGATGCTCGTCACCAATATGCTTCTGTGTAAGCGCAT 531
Oy      68 ttgcacaatgcgtttaagaatacaacttgyaactaaagccgtaaatgtcgtlatctgaagaat 127
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      532 TGGCGTAGAGCATTTAAACCAACAAATGGGTCCAAGCGGAGTAAAGTTGTTTAAAGAAAT 591
Oy      128 ctttgtgtccccagttatatacaaaagttgtgtatctgtgtgcaaaaaaataatgtaactg 187
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      592 CATTTGGTGCACCTACTGTAATAAGACGGGTCTTCTTGCAANAAGAAATGAATTCG 651
Oy      188 aagalaagttlgaanaataatggcgctcaaatgttlaaagaatagctcccacaacatgcg 247
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      652 AAAACCGTTTCAAANAATATGGCGCACCAATGCTTAAGAAATTCGTGTAAACACTTCG 711
Oy      248 atattgtcgtgtgaatgaactacaagaacagtcctctgcacaagctattatcgtgaag 307
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      712 ATACCGCAGAGTGATGTGCTACTACTGTGTAACGGTATTTGGACAGCTGCATTTGCTGTGAAG 771
Oy      308 ggtgtaaacactgtgaacagctgtgtglatctcattgtgccaattaacgttgcatagataag 367
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      772 GCCACAAAGCAGTTCCACACGCCATGAACCCCAATGCATTTGAACGAGGATATTGATAAG 831
Oy      368 clgtgtgtcgtgttctlaaagaactaaagcgacatcaaacgctactcgtgcaccaaaag 427
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      892 CAGTTGCCGCGCASTTACCAAAAATTTNCAAGAAATGCTCTTAAGCTTGCACAAAGACGGTAAG 891
Oy      428 aatatgtctcaagtgtgaacacattcttcgaabaactctgatacaacaatagtlaatatcag 487

```

D	892	CAATTGGACAAGTAGGACTAATTTCTGCGAAGCTCTATPAAAGCAATCGGTTCTATCATTTG	951
O	488	ctgaagaatctgagctaaagcttggaaagaagctgttaacaaagcttggagaagctaaagtc	547
D	952	CTGAAGCAATGCGAAAAAGTAGTAAGAAGGCTGTTATTACTAGAGAAGTGGTAACAGCC	1011
O	548	ctgaacacacatagatctgtgtctgaaagaaagagcttgacagctgtcactccctccat	607
D	1012	TTGAAACGAAATTTGCTGTTGGTAAGGAGTATGACAGTTTATCGTGTTCATTTCTCCAT	1071
O	608	actctgtaacataactccgagaanaatggtctgtgaacttgataacctatacccttatccctt	667
D	1072	ACTTCATCAACAACTACCAAAACATGACGTCTAGCTTAAGCATCCTTTTCATCTCTTGG	1131
O	668	atggaagaaagatattacatagcatgaaagaatgcttacaactctagaagaagcttgctaaag	727
D	1132	TTGACAAAGAAATTTTCAACAATTCTGACATGTTGCTGTTATTGGAAGACAGTTGCCAAT	1191
O	728	taaacgcgtccactccctctattctgtctgaaagcgtagaagaggtaagcaacttgcaactgt	787
D	1192	CAGAGCCGCTTTATTGATCATTTGTAAGATGTTGAAGTGAAGCATTAAGCAACTGTGG	1251
O	788	tagtcaataagcttcgtgtgagagcaactccaaagtctagccgttaaaagctctgtttctgt	847
D	1252	TTGTTAACAAACATCCGGCGATTGTTAAATGTCGCCGGTTAAAGCACCCTGTTTCGGTG	1311
O	848	aacgcgtlaaagctatctgtgaaagatatctctccctctctgaaagagaagcaatatctg	907
D	1312	ATCGTCCTAAAGCCATTTTCAAGACATTTGCTATCTTACTGCGGGCAAGCATTTTCAG	1371
O	908	aagatctgtgtataaagcttgaagaaatgtaagctctgtctctctcttgaagaacagctaaacgt	967
D	1372	AAGAAGGGGAGTACTAGTTTGAAGAAGTCTACTTTAACAGCCTAGGCACATGCAAAACCG	1431
O	968	tagtatttgacaagaanaataactactatacgtgtgaatggtgtcgtggaanaatcagaagata	1027
D	1432	TTTGCTGTACCAAAAGAAACCCACACATCTTTACCGCTGAAGCTAAAGCGCTGCATCA	1491
O	1028	aagctcagatctaaacaactctgtgcacaacttgaagaagaaacagagctcagatctgtctg	1087
D	1492	ATCGCGCTATTACTCAAAATTCTGTCCTCAAAATGGAAACAAACACTTCTGATTACGATCCG	1551
O	1088	aaaactctcaagaacgctctgtcacaactctgtgtgaaatgaactgttacaactgtctgag	1147
D	1552	AAAATATTACAAGACCGCGTAGTAATTTACTCTGGCGGTGTGGCGTCAT -TAAGTTGGTG	1611
O	1148	ctgtactctgaacactgaaatgaaagagaagaagatcgtgtagaagaatgtctctcaatgcaa	1207
D	1611	CTGTGACTCCGAATC----GAATGAAGAGAAACCTGTGTAGAACGCCCTTTCATATGCAA	1665
O	1208	caagagcgtccgagcttgaagaagagatgtccctcgtgtgtgtgtgaactgtcttctgcgcctca	1267
D	1666	CTCTGCTGCGCAGTAGAAGAAAGGCAATTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1722
O	1268	ttaaaagctctgatatataaactctgtctgatatgatagaactctgtgaactataatca	1327
D	1726	AAAAAGCATTTAGATGGCTGTAAAGAGGCAAAATCTGACCAAGACATGGGTATTAACTATCC	1785
O	1328	tcgctgtctctctgaaagagcctttaaagctaaatgtgtcgaatgtcgtgataatgaaagt	1387
D	1786	TACGTCGCGGCATGCAATCTCTCTCTCGGCCAAATCTGTACCAATGACAGGTATTGAGCCCT	1845
O	1388	ctatctgtgtagaanaaagcttcgtgaagcaaaaagaatggttttgatttaatgtctgaatga	1444
D	1846	CTGTATTATGTTAAACAAGATGCTGAAGCAAAACATTAATTTGGTTTCAACGACGATCTAG	1905
O	1448	gagataatgaaagacattataaagctgtgtgtcatctgatatcctaaagaagctaaagctatgt	1507
D	1906	GGCAATTAAGGAGACATGTTGAGATGGGATTTCTTATCTCCAACTAAAGATGCTGTACG	1965
O	1508	cattacaataatgcacatcagtaagctcctctactcttcaactacagaatgtgcatactgct	1567

QY 1081 gatcglgaaaaactlcaagaagcgtctgcaaaactgttgtagagtagctgtatccat 1140
DB 1748 GATCTCTAGAAGCTGCAGAGAAAGATGGCTTAACCTTCTGCTGCTGCCGCTGATATAA 1807
QY 1141 gttagagctgctactgaagaatgaagaagaagagctgtagaagaatgctccta 1200
DB 1808 GTTGTGCTCTTTACAGAGCTGAGATGAAGAGAGAAAGATCGTGTGATGATGCTTTA 1867
QY 1201 aatgcacaagaagctgcgtgtgagaagaagtagctgcccgtgtgtgtactgtcttc 1260
DB 1868 CACTCTACTGTCGCGCTGAGAGAAAGATATCTTCTGCTGCTGCTGCTGCTTTATTT 1927
QY 1261 cgcctccatlaaagctcgtgtgtatataaagcgtgtagatgaatgaactgtcgtgact 1320
DB 1928 AGAGCACAAAAGCATTAGATGGCTTACAGGTGAATAATGATCAAAACCATGATATA 1987
QY 1321 aatacatcgcctgctcctctgaagaagccttctacgtcaaatgctgtcaaatgctgctat 1380
DB 1988 GCGCTACTTGAAGAAAGCAATAGAAAGCTCTTAAGACAGATGATCAAAATGCTGGCGGT 2047
QY 1381 gaaggtctcatgtgtgtagaagaagctgtgaaaccaaagaatgtgtttagatctaatgtc 1440
DB 2048 GAGCTTTCTAGTTGTTAAACCAAGTTAAAGCTTAATCAAGCTAATGCTTATATAAGCT 2107
QY 1441 gcataagaagaatagagaagccttattaaagcgtgtgtcatgtatcctaaagaagtaca 1500
DB 2108 GCAAAATGATACTTATGGTGTATGGTGTGAGATGGGTATTTTATGCTTACTTAAGTTACT 2167
QY 1501 cgtatgtctactcaaaatgagacatcagtagcctcctactctcaactacagaatgctgc 1560
DB 2168 CGTTACACTTACACACTGTGTGCTTCAATGCTGACTTATGATCACTACAGAGCGATG 2227
QY 1561 attgcgtgaaaaacagaacctaagaagaatgctcctatgcccgtgaggtgtgtgtgt 1620
DB 2228 ATCGCTGACGATTAAGCAAGCTGCTCTGCTATGCTGCTATG---GCCGCTGGCATGGCGGT 2284
QY 1621 atgggtgtgtatg 1632
DB 2285 ATGCTTGCAATG 2296

RESULT 14
AX067460
LOCUS AX067460 96109 bp DNA PAT 24-JAN-2001
DEFINITION Sequence 35 from Patent WO0078968.
ACCESSION AX067460
VERSION AX067460.1 GI:12545080
KEYWORDS
SOURCE Moraxella catarrhalis.
ORGANISM Moraxella catarrhalis.
Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
Moraxella.
:JENCE 1 (bases 1 to 96109)
:JTHORS Lagace,R.E., Patterson,C. and Berg,K.L.
TITLE Nucleotide sequences of moraxella catarrhalis genome
JOURNAL Patent: WO 0078968-A 35 28-DEC-2000;
Incyte Genomics, Inc. (US)
FEATURES
source 1. 96109
/organism="Moraxella catarrhalis"
/db_xref="taxon:480"
BASE COUNT 28783 a 18910 c 20341 g 28075 t
ORIGIN

Query Match 45.0%; Score 741.8; DB 10; Length 96109;
Best Local Similarity 66.3%; Pred. No. 2,4e-119;
Matches 1086; Conservative 0; Mismatches 547; Indels 6; Gaps 1;

QY 8 ctaagaagaatccttctgtagctaaagccgtgaaaaacttccagaggtgtatagataaac 67
DB 49280 CAAAGAATGTAGTTTGGCTCAAAATGCGCCGAGAAAATGATGATGCTGTTAAACATT 49339

QY 68 ttgcaaatgctgttaaaagtaaacacttggaccctaaagccgttaatgtcgtatattgaaagt 127
DB 49340 TGGCGGATGGCTTCAAGCTTACCTTACCTTCAAAAGGCCCAAGTGTGATTAAT 49399
QY 128 ctttggctcccccagttattacaagaatgtgtatcgtgtctgcaagaagaatgaacttg 187
DB 49400 CATTGGTGGCCGACCATCAACCAAGATGCGGTGATGTTGCCAAGAGATTAACTTG 49459
QY 188 aagataagtttgaataatagtggtcgtcaaaatggtttaaagaagttagctcccaaaactagcg 247
DB 49460 AAGATAATTTGAAACATGGGTGCACAAATGGTGTGCAAGTTGCTTAAGCCCAATG 49519
QY 248 atattgtgtgtgtgtagaacctacaagaagcagccttgcacaagaagctattatcgtgaag 307
DB 49520 ATGTGGCAGCGGTGAGCAACAACCCGACAGTACTTGTCTCAAGCCCATCTTGGTTAGG 49579
QY 308 gtgtaaactgttagcagcgtgtcgtatccctaaagtcataaagcgtgcatagataaag 367
DB 49580 GCATGAANAAGCGTTTGTGCGAGCATGATCATGATGATCTTAAGCCGTGATGATTAAG 49639
QY 368 ctgtgtgtgtgttaactaaagaactaagcgacattacaagaagcctcgtgaccaaaag 427
DB 49640 CCGTACCGCGCGCGTGTGAAGAAATCCGTCTATCTTCAGACACCTGCTAATGACCATTAAG 49699
QY 428 aatatagctcaagttggagaccatttctgcaaacctgtatgataaagaatagtaatactag 487
DB 49700 CGATTGCTCAAGTTGGCTCAATTTTCAGCAAACTGATGCTACCATTTGTAGCTTAATCT 49759
QY 488 ctgaagctatggtctaaagcttggaaaagaggtgttaccacagttgaaagaactaaagtc 547
DB 49760 CCAAGGCAATGGAAGACAGTTGGCAACAGCGCTTATCACAGTTGAAGAGGTTCAAGTT 49819
QY 548 ttgaactatcatatagatgtgtgttgaagaatgaagttgacggtgtactcctccat 607
DB 49820 TTGAAGATGGCGTTGAAGTGTGAGGCGCATGAGTTGACCGGCTATATACGCCCTT 49879
QY 608 acuttgtaactaatcccgaaaatagttgttgtagacttgtaacctatacttctgtga 667
DB 49880 ACTTGGCAATTAACAAGACAGCTTGACTTGTGATTTGACAAATCTTAATTTCTTTTGC 49939
QY 668 atgagaaaagaatctactagatgaaagacatgctacccaacttgaagaagaagtgtctaag 727
DB 49940 TCGATTAATAAATCTCAAAATATTGCTGAGATTTGGCCACTACTGAAAAGTGATGCAA 49999
QY 728 taaacgtcactccttattatgtgtgaaagcgttagaaggtgaaagcactgtcaactgtg 787
DB 50000 CCAAGCCCTCCGCTAATTAATCTGCTGAAGATGTTGAATAAGCAATTGCGCAACTTGC 50059
QY 788 tagtcaatgaagctcccggtgagagcactccaagttgtagccgtgaaagcctcgtgtttgtgt 847
DB 50060 TTGTCAATACTTTCGTGGCGGATTAATAAACTTGTGGGTTAAAGCCCAAGTTTGTGTG 50119
QY 848 aagcgcgtlaaagctatgcttgaagatactctactcactctgagaggaagcaataattgt 907
DB 50120 ATGCGCGTAAAGCATGCTTCAAGACATTTGCAATTTAAACAGTGGCGGTGATTATCAG 50179
QY 908 aagatcgtgtatataaagcttgaagaatgaagctgtcttctcttaagaaacagctaaacgtg 967
DB 50180 AAGAAATGCGCTTAAGCTTGAAGCTGTGAGACTGCTGAGATTGAGCAATTTGGTATCGCCAAAAG 50239
QY 968 tagtattgcaagaagaataactactatcgtgttgaaggtgtgagaaatcgaagaatactta 1027
DB 50240 TAACCATTTGTAAGAAAGAACCCGCTCATTTGTCATGGTCAAGATTAAGCCACGCAATTG 50299
QY 1028 aagctcgtatlaaacaactcgtgtcacaaatltgaagaagaacagctcagatcatgtatcgtg 1087
DB 50300 AGGCGGTGTAAGATCAATTCATGCTGCTGAGTGAAGATCAATTCATGATACGACAAAG 50359
QY 1088 aaaaacttcaagaacgtcttgcacaactgtgtgtgtagagtagcgttatacattgtgtgag 1147
DB 50360 AAAAATTTCAAGACCGGTGTGCAAAACTATACAGCGGTGTGTCAGTCAATCAAGTGGCGTG 50419
QY 1148 ctgctactgaaactgaaatgaaagaagaagatcgtgtgaagaatgctcttaaatgcaa 1207

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2001, 15:39:34 ; Search time 136.77 Seconds

(Without alignments)
7561.263 Million cell updates/sec

Title: US-09-077-574A-1

Perfect score: 1647

Sequence: 1 atgctcttaagaataatcct.....gtatgacgtatgtactag 1647

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDSI/gcgdata/geneseq/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/NA2001.DAT:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1647	100.0	1647	18 AAT69201	Lawsonia intracellularis
2	1647	100.0	1647	18 AAT69203	Lawsonia intracellularis
3	741.8	45.0	96109	22 AAT68348	Genomic fragment #
4	713.6	43.3	2465	12 AAQ13136	Hyp operon. Chlam
5	710.4	43.1	2223	12 AAQ13137	Hyp operon. Chlam
6	704.2	42.8	1661	20 AAX86155	DNA encoding a Str
7	698.4	42.4	3625	20 AAX12979	Enterococcus faeca
8	689.8	41.9	1654	20 AAX86153	DNA encoding a Str
9	689.8	41.9	1926	22 AAT55036	Nucleotide sequence
10	687.6	41.7	5365	19 AAV52210	Streptococcus pneu
11	686	41.7	1038602	20 AA201425	Complete genome se

12	684.4	41.6	1647	22 AAC90805	Escherichia coli C
13	682.8	41.5	1647	21 AAA48498	E. coli groEL codi
14	682.8	41.5	4524	21 AAA48500	Escherichia coli g
15	652.8	39.6	1665	20 AAX86152	DNA encoding a Str
16	644.8	39.1	1662	20 AAX86154	DNA encoding a Str
17	642.4	39.0	1838	14 AAQ47926	hsp60 DNA. Helico
18	642.4	39.0	1838	14 AAQ47924	Hsp gene. Helico
19	639.2	38.8	1724	19 AAX14495	H. pylori GHP0 118
20	632	38.4	1635	21 AAA13000	Neisseria meningit
21	632	38.4	1635	21 AA254508	Neisseria meningit
22	632	38.4	1635	21 AA254509	Neisseria meningit
23	632	38.4	13423	21 AA481527	N. meningitidis pa
24	632	38.4	349980	21 AAT21612	Neisseria meningit
25	632	38.4	837096	21 AA481489	N. meningitidis pa
26	610.4	37.1	2284	16 AA090181	Helicobacter pylor
27	610.4	37.1	2284	17 AAT45681	H. pylori heat sho
28	610.4	37.1	2332	16 AAT75321	Heat shock protein
29	609.6	37.0	910715	20 AA220248	Heat shock protein
30	598.4	36.3	1635	21 AA254507	Borrelia burgdorfe
31	550	33.4	1847	22 AA545450	Neisseria gonorrhe
32	548.2	33.3	1725	22 AAT32452	Pseudomonas aerugi
33	512.8	31.1	2242	18 AAT58403	Human heat shock p
34	501	30.4	1925	21 AAC37621	Arabidopsis thalia
35	495.4	30.1	1909	21 AAC49855	Arabidopsis thalia
36	485.2	29.5	2193	11 AAQ04659	Heat shock protein
37	471	28.6	2127	21 AA446168	B. pseudomallei Gr
38	466.6	28.3	2059	21 AAC44547	Zea mays DNA fragm
39	457.2	27.8	1647	17 AAT14265	Brevibacterium fla
40	454	27.6	1986	21 AAC44753	Zea mays DNA fragm
41	447.6	27.2	1996	21 AAC38636	Arabidopsis thalia
42	446	27.1	1953	21 AAC50118	Arabidopsis thalia
43	445	27.1	1994	21 AAC50121	Arabidopsis thalia
44	445.6	27.1	3613	9 AAN80339	Clone Y1178 insert
45	441.2	26.8	1777	19 AAV59425	Heat shock protein

ALIGNMENTS

RESULT 1	
AAT69201	
ID AAT69201 standard; DNA; 1647 BP.	
XX	
AC AAT69201:	
XX	
DT 20-AUG-1997 (first entry)	
XX	
DE Lawsonia intracellularis GroEL DNA.	
XX	
KW Intestinal disease; porcine proliferative enteropathy; vaccine;	
KW GroEL, heat shock protein; ss.	
KW	
OS Lawsonia intracellularis.	
XX	
PN WO9720050-A1.	
XX	
PD 05-JUN-1997.	
XX	
PF 29-NOV-1996; 96KO-A000767.	
XX	
PR 30-NOV-1995; 95AU-0006911.	
PR 30-NOV-1995; 95AU-0006910.	
XX	
PA (DARA-) DARATECH PTY LTD.	
PA (PIGR-) PIG RES & DEV CORP.	
XX	
PI Hasse D, Panaccio M;	
XX	
DR WPI: 1997-310605/28.	
DR P-PSDB: AAM16678.	
XX	
PT vaccine for treating or preventing Lawsonia intracellularis	
infection - especially in pigs, containing non-pathogenic form of	

pr bacterium or its components
 xx
 ps Claim 12, Page 38-42, 94pp; English.
 cc A DNA molecule (AAT69201) codes for the GroEL heat shock protein
 cc (AAH16678) of Lawsonia intracellularis, the causative agent of porcine
 cc proliferative enteropathy (PPE). A genomic library was prepd. from
 cc L. intracellularis PPE lesion isolates and screened with rabbit
 cc anti-L. intracellularis antiserum. Phagemid DNA was isolated from
 cc individual clones and sequenced. GroEL and GroES (see also AAT69202)
 cc sequences were identified and another 13 clones were subsequently
 cc (AAT69203-15) isolated. These nucleic acids can be used to produce
 cc recombinant polypeptides useful in vaccines against intestinal
 cc diseases such as PPE, including recombinant vaccines utilizing
 cc bacterial, fungal or viral vectors. They can also be used as
 cc genetic vaccines and in diagnostic assays.
 xx
 sq Sequence 1647 BP; 536 A; 277 C; 368 G; 466 T; 0 other;

Query Match 100.0%; Score 1647; DB 18; Length 1647;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1647; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 atggctctaaagaatccttcttgatgctaaagccggtgaaaccttaccaggtgta 60
 Db 1 atggctctaaagaatccttcttgatgctaaagccggtgaaaccttaccaggtgta 60
 QY 61 gataaacttgcgaatgctgttgaagaacacttgcgaactaaagccgttaagtcttatt 120
 Db 61 gataaacttgcgaatgctgttgaagaacacttgcgaactaaagccgttaagtcttatt 120
 QY 121 gaaaagcttcttggttcccaagtattacaagaagatggttactctgtgcaaaagaat 180
 Db 121 gaaaagcttcttggttcccaagtattacaagaagatggttactctgtgcaaaagaat 180
 QY 181 gaacttgcgaagaatgcttgaagaatattggtggtcgaactaaagttaaaagaagctcccaa 240
 Db 181 gaacttgcgaagaatgcttgaagaatattggtggtcgaactaaagttaaaagaagctcccaa 240
 QY 241 actagcgatattgctgtgtgaaggactacacagcaacagcttctgcacagctattat 300
 Db 241 actagcgatattgctgtgtgaaggactacacagcaacagcttctgcacagctattat 300
 QY 301 cgtgaaggtgttaaaacttgaagcagctgtgcgttaactctatggtccattaaagctggcata 360
 Db 301 cgtgaaggtgttaaaacttgaagcagctgtgcgttaactctatggtccattaaagctggcata 360
 QY 361 gataaagctgttctgtcttactaaagaactgaagcgacattacaagcttactctgac 420
 Db 361 gataaagctgttctgtcttactaaagaactgaagcgacattacaagcttactctgac 420
 QY 421 caaaaagaatagcgaacttgaagaacttctgcgaacttgcatacaaaactaaggttaat 480
 Db 421 caaaaagaatagcgaacttgaagaacttctgcgaacttgcatacaaaactaaggttaat 480
 QY 481 atcatgctgaagctatgctgaaggttgaagaaggtgttaccagtctggaagct 540
 Db 481 atcatgctgaagctatgctgaaggttgaagaaggtgttaccagtctggaagct 540
 QY 541 aaaggtctgaagaactatagatgtgtgtgaagaaggtgaagcttgaacgtgtctacac 600
 Db 541 aaaggtctgaagaactatagatgtgtgtgaagaaggtgaagcttgaacgtgtctacac 600
 QY 601 tctccatacttctgaactaaactctggaagaatgttctgtgaacttgaacttatac 660
 Db 601 tctccatacttctgaactaaactctggaagaatgttctgtgaacttgaacttatac 660
 QY 661 ctctgttaatgaagaagaattactagcatgaagaagacatgttaccatcttaagaagct 720
 Db 661 ctctgttaatgaagaagaattactagcatgaagaagacatgttaccatcttaagaagct 720
 QY 721 gctaaagtaaacgctccactcttatattgtcgaagagctagaaggtgaagacttgc 780

Db 721 gctaaagtaaacgctccactcttatattgtcgaagagctagaaggtgaagacttgc 780
 QY 781 acactgtgaactaataagctccgtgagacgtcccaagtgttaacgttaaaagctctgt 840
 Db 781 acactgtgaactaataagctccgtgagacgtcccaagtgttaacgttaaaagctctgt 840
 QY 841 ttgtgtgaagcgcgttaagactatgtctgaagatactgtctactctggaagaaga 900
 Db 841 ttgtgtgaagcgcgttaagactatgtctgaagatactgtctactctggaagaaga 900
 QY 901 atattggaatgtgtgttaaaagcttgaagaagtaagcttgccttcttaagaagaagct 960
 Db 901 atattggaatgtgtgttaaaagcttgaagaagtaagcttgccttcttaagaagaagct 960
 QY 961 aaacgtgaacttattcaagaagaataactactactcgtgatgtctggaagaatcga 1020
 Db 961 aaacgtgaacttattcaagaagaataactactactcgtgatgtctggaagaatcga 1020
 QY 1021 gatattaagctcgaacttaacaactcgtgtcacaactgaagaagaacagctcagattat 1080
 Db 1021 gatattaagctcgaacttaacaactcgtgtcacaactgaagaagaacagctcagattat 1080
 QY 1081 gatcgtgaaaacttcaagaagctcttgcgaactgtgtgtgtgagtagtcttaccat 1140
 Db 1081 gatcgtgaaaacttcaagaagctcttgcgaactgtgtgtgtgagtagtcttaccat 1140
 QY 1141 gttggaagctgtactgaactgaactgaagaagaagaagatcgttgaagaagtgtccta 1200
 Db 1141 gttggaagctgtactgaactgaactgaagaagaagaagatcgttgaagaagtgtccta 1200
 QY 1201 aatgcaacaagagctcgtgttgaagaagatattgtccctgtgtgtgtgtgtgtgt 1260
 Db 1201 aatgcaacaagagctcgtgttgaagaagatattgtccctgtgtgtgtgtgtgtgt 1260
 QY 1261 cgtctccattaaagctcttgaatgaatgaactcgtgtagatgaatgaactgttgaact 1320
 Db 1261 cgtctccattaaagctcttgaatgaatgaactcgtgtagatgaatgaactgttgaact 1320
 QY 1321 aatatacgcgtctctcttgaagaagcttgaactcaaatlgtcgaactgtcgtcatt 1380
 Db 1321 aatatacgcgtctctcttgaagaagcttgaactcaaatlgtcgaactgtcgtcatt 1380
 QY 1381 gaaagcttcatgttctgaagaaggtcgtgaacccaagaagatggttctgatttaactgt 1440
 Db 1381 gaaagcttcatgttctgaagaaggtcgtgaacccaagaagatggttctgatttaactgt 1440
 QY 1441 gcatcagaagaatgaagaacttataagctgtgtgtcatgtatgcttaaaagaattaca 1500
 Db 1441 gcatcagaagaatgaagaacttataagctgtgtgtcatgtatgcttaaaagaattaca 1500
 QY 1501 cgtattgcattacaagaatcgaagatcagtagccttcttaacttaactacagaatgcgt 1560
 Db 1501 cgtattgcattacaagaatcgaagatcagtagccttcttaacttaactacagaatgcgt 1560
 QY 1561 atgtcgtgaagaacacgaacttaaaagaatattcctcgtcgtgtgtgtgtgtgtgt 1620
 Db 1561 atgtcgtgaagaacacgaacttaaaagaatattcctcgtcgtgtgtgtgtgtgtgt 1620
 QY 1621 atgggtgtatgaagcgtatgtactag 1647
 Db 1621 atgggtgtatgaagcgtatgtactag 1647

RESULT 2
 AAT69203
 ID AAT69203 standard; DNA; 4972 BP.
 xx
 AC AAT69203;
 xx
 DT 20-AUG-1997 (first entry)
 xx
 DE Lawsonia intracellularis vaccine candidate DNA.

XX Intestinal disease; porcine proliferative enteropathy; vaccine; ss.
 KW Lawsonia intracellularis.
 XX
 OS
 XX
 PN W09720050-A1.
 PD
 XX 05-JUN-1997.
 PF 29-NOV-1996; 96MO-AU00767.
 XX
 PR 30-NOV-1995; 95AU-0006911.
 PR 30-NOV-1995; 95AU-0006910.
 XX
 PA (DARA-) DARATECH PVTY LTD.
 (PIGR-) PIG RES & DEV CORP.
 PA
 AA Inasse D, Panaccio M;
 DR WPI: 1997-310605/28.
 XX
 PT Vaccine for treating or preventing Lawsonia intracellularis
 PT infection - especially in pigs, containing non-pathogenic form of
 PT bacterium or its components
 XX
 PS Claim 14; Page 47-50; 94pp: English.
 XX
 CC DNA molecules (AT69203-15) were isolated from an immunoscreening of
 CC a Lawsonia intracellularis library using experimental sera from
 CC vaccinated pigs.
 CC And GroES (AT69201-02) sequences, encode putative vaccine candidates
 CC (see also AW16678-85) useful for protection of animals and birds
 CC against intestinal diseases, esp. protection of pigs against
 CC porcine proliferative enteropathy (PPE). They can also be used as
 CC genetic vaccines.
 CC
 XX
 SQ Sequence 4972 BP; 1717 A; 871 C; 831 G; 1532 T; 21 other:
 Query Match: 100.0%; Score 1647; DB 18; Length 4972;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1647; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 atgctctcaagaatcctcttctgactgaagcccgtaagaactcttcagagtgta 60
 2048 atgctctcaagaatcctcttctgactgaagcccgtaagaactcttcagagtgta 2107
 QY 61 gataaacttgcaaatgctgttaaaagtaaacacttggaacttaaggccgtaatgctgta 120
 DB 2108 gataaacttgcaaatgctgttaaaagtaaacacttggaacttaaggccgtaatgctgta 2167
 QY 121 gaaaagcttctgttcccggtatatacaaaagatggtatctgttgcagaaagaatt 180
 DB 2168 gaaaagcttctgttcccggtatatacaaaagatggtatctgttgcagaaagaatt 2227
 QY 181 gaactlgaagaataagtttgaataatagcgctcaaatggtttaaagaatgctcccaaa 240
 DB 2228 gaactlgaagaataagtttgaataatagcgctcaaatggtttaaagaatgctcccaaa 2287
 QY 241 actagcgaatctgtcgtgtaagtaactaacacagcaacagtccttgcacaaagctatatt 300
 DB 2288 actagcgaatctgtcgtgtaagtaactaacacagcaacagtccttgcacaaagctatatt 2347
 QY 301 cgttaaggtgttaaaactgtgtgacgctgctgtatctatactgagccatlaaacgctgata 360
 DB 2348 cgttaaggtgttaaaactgtgtgacgctgctgtatctatactgagccatlaaacgctgata 2407
 QY 361 gataaagctgtgtgtcgttactcaaaagaaactaaagcgaatctataaagccctaccgttac 420
 DB 2408 gataaagctgtgtgtcgttactcaaaagaaactaaagcgaatctataaagccctaccgttac 2467
 QY 421 caaaagaataatgctcaagtgtgaacattctgcacaaactgtatatacaaatagtaatt 480

DB 2468 caaaagaataatgctcaagtgtgaacattctgcacaaactgtatatacaaatagtaatt 2527
 QY 481 atcataagctgaagctatagctcaaaagtgtgaagaagaggtgttatcacagtttgaagaact 540
 DB 2528 atcataagctgaagctatagctcaaaagtgtgaagaagaggtgttatcacagtttgaagaact 2587
 QY 541 aaaggtctgaactacatactagatgtgttgaagaagaaacttgacagctggtaccc 600
 DB 2588 aaaggtctgaactacatactagatgtgttgaagaagaaacttgacagctggtaccc 2647
 QY 601 tcccaatacttctgaactaactccctgagaanaatgtgttctgaactgtatacccttatac 660
 DB 2648 tcccaatacttctgaactaactccctgagaanaatgtgttctgaactgtatacccttatac 2707
 QY 661 ctgttgataggaanaagatctactagactgtgaagaacatgctaccacaaactttgaacaaatt 720
 DB 2708 ctgttgataggaanaagatctactagactgtgaagaacatgctaccacaaactttgaacaaatt 2767
 QY 721 gctaaagtaaacctccacactccttatactgtctgaagcgttagaaggtgaagcacttga 780
 DB 2768 gctaaagtaaacctccacactccttatactgtctgaagcgttagaaggtgaagcacttga 2827
 QY 781 acaacttgatgaataaagctccgtgagacactccaagtgttagccgttaaaagctcctgt 840
 DB 2828 acaacttgatgaataaagctccgtgagacactccaagtgttagccgttaaaagctcctgt 2887
 QY 841 ttgtgtgaacgcgttaaaagctatgctgtgaagaatattgctacactctactgagaggaagca 900
 DB 2888 ttgtgtgaacgcgttaaaagctatgctgtgaagaatattgctacactctactgagaggaagca 2947
 QY 901 atatttgaagatcgtgtgataaagcttgaagaatattgctacactctactgagaggaagca 960
 DB 2948 atatttgaagatcgtgtgataaagcttgaagaatattgctacactctactgagaggaagca 3007
 QY 961 aaacgtgtgattatgacaagaataactactactcgtgtgaatgtgtgtcgtgaataacagaa 1020
 DB 3008 aaacgtgtgattatgacaagaataactactactcgtgtgaatgtgtgtcgtgaataacagaa 3067
 QY 1021 gataatgaagctgagtttaaaacaaattcgtgtcacaaatttgaagaacaaagctcgtatatt 1080
 DB 3068 gataatgaagctgagtttaaaacaaattcgtgtcacaaatttgaagaacaaagctcgtatatt 3127
 QY 1081 gatcgtgaanaaacttcaagaacgctcttgcacaaactgtgttgcagtaagcgttatccat 1140
 DB 3128 gatcgtgaanaaacttcaagaacgctcttgcacaaactgtgttgcagtaagcgttatccat 3187
 QY 1141 gctggaagctgctactgaaactgaaatgaaagagaagaagatcgctgtaagaatgctcta 1200
 DB 3188 gctggaagctgctactgaaactgaaatgaaagagaagaagatcgctgtaagaatgctcta 3247
 QY 1201 aatgcaaaagagctgcgtgtgaagaagaagtatttcctctgtgtgtgtgtaagcgtttgtc 1260
 DB 3248 aatgcaaaagagctgcgtgtgaagaagaagtatttcctctgtgtgtgtgtaagcgtttgtc 3307
 QY 1261 cgcctcattaaagctcgtatgatatataaactctgatatgtaactgtcgtgaact 1320
 DB 3308 cgcctcattaaagctcgtatgatatataaactctgatatgtaactgtcgtgaact 3367
 QY 1321 aatatcatcgctgcttctctctgaaagcctttacgccaactgctgcgaatgctgctat 1380
 DB 3368 aatatcatcgctgcttctctctgaaagcctttacgccaactgctgcgaatgctgctat 3427
 QY 1381 gaaaggtctatgtgtgtgaagaagaagtctgtgaacaaagaagtgtttgtgattaaagt 1440
 DB 3428 gaaaggtctatgtgtgtgaagaagaagtctgtgtgaacaaagaagtgtttgtgattaaagt 3487
 QY 1441 gcatcagagagataatgaagacattataaagctgtgtcatgtatcctaaagaagttata 1500
 DB 3488 gcatcagagagataatgaagacattataaagctgtgtcatgtatcctaaagaagttata 3547
 QY 1501 cgtatgtcatatacaaaatgacatcagtagcctccttaacttcaactaagaatgctc 1560
 DB 3548 cgtatgtcatatacaaaatgacatcagtagcctccttaacttcaactaagaatgctc 3607

[illegible]

Db	49400	catcttgcgctgcgcgaccatccacaaagatggcgctgacatctgccaaggaattgtaacttg	49455
Qy	188	aagataagtttgaanaatactggtgcgtcaaatggttaaagaagtgcgtcccaaatatgctg	247
Db	49460	aagatlaaatcttgaanaacatggtgtgcacaaattgctgtcgtgaagttgctcttaagaacaa	49519
Qy	248	atactgcgtgtatgtgaacttaacagacaaagaactctgtccagaagatttatctcgtgaag	307
Db	49520	atgtgtgcaggtcgtatgtgaacaacacacccgcgaacagttactgtctcaagccattgtgtgaag	49579
Qy	308	gtgtlaaacctctgtagagcgtgtgtcgtatccctatgctgccaataagctggcatagataaag	367
Db	49580	gcattgaaaacggtttgctgtcagtcgatccaatgtgatactaaagcgttggttatgttaaaag	49639
Qy	368	ctgtgtgtcgttactataagaactaagcgtacatactcaaaagcttactcgtgtcaaaaag	427
Db	49640	ccgtlaacgcgcgcctgtcttgtagaagaatccgtgtctatctcgacactcgtctaatgtccaataag	49699
Qy	428	aaataacctaagtttggaacactcttcctgcgaactctgtataccaacaataggtaatatcaatag	487
Db	49700	cgattgtctaaattgtgtctcaattcttcgcgaactcgaatgtctccattgtgttggtatact	49755
Qy	488	ctgaagctatgtgcctaaagtttggaaaaagtgagttgtatccacagttgaggaagctaaagtc	547
Db	49760	ccaaggtcgaatgtgaaaacaagttctggcaaaacaagcgttatccacagttgaggaaggtgtcaggt	49819
Qy	548	ttgaaactcatatagatgtgtgtcttggaaggaatgaagtgtgacggtgtgctacccctccat	607
Db	49820	ttgaaagatgcgtcttgaaagttaagtttggaaggtacgtttagacgtgtgtataccaaacctt	49879
Qy	608	actctgtlaactlaactcctgtgaanaaatggtttgttgaaacttgatatacccttatcccttgta	667
Db	49880	actctgcgaataaacaagcagcagctgtgactgtgaaatttgacaatccattatcttttttg	49939
Qy	668	atgagagaaaagtttctgctgcgtgaagaagacatgctaccaatctttagaacagtgtctaag	727
Db	49940	tcgataaaaaaaactccaacatctcgtgaatctgtgcactactgtgaaaaagtgatgtgcaaa	49999
Qy	728	taaacgctccacatcccttatctatctgtcgtgaagaagctgaagaagtgaaagcaactgtcaacactg	787
Db	50000	ccagcgcgtccgttatatactatctgtcgtgaagaatgttgaaaaatgaagcatgtgcacaacttgg	50055
Qy	788	tagtcaataagcttccgtgtgagcactcccaagttgttagccgttaaaagctccctgtgtttggtg	847
Db	50060	ttgtcaatctcttctgtgtgtgtgcggtatataaaactctgtgcgtttaaagcgcgaagttttgtg	50119
Qy	848	aacgcgctaaagctctgtcttggaagatattgtctatcccttactgtggaagagaagaacatattg	907
Db	50120	atcgcgcgtaaagcctatgtcttcaagacatctgtcaacatcttaacaaagtgtgcgtttatccag	50179
Qy	908	aagatcgtgtgataaagcttggaanaatgtgaagctgttctctctttaggaacagctaaacgtg	967
Db	50180	aagaagttggtcttaagcctctgtgactgtcgtgaatttgagcatttgggtttactgccaanaag	50239
Qy	968	tagttattgtgacaaaataaatactactatctggttgaatgtgtgtggaataatcagaagatatla	1027
Db	50240	taacacctgtgtataagaanaaacaccgtcatctgttgtagtgtgcaggtgataaagccagcatgtg	50299
Qy	1028	aagctctgaagttlaaacaatactctgtgcacaatactgaagaagaacaagcctcagatatctgctgtg	1087
Db	50300	aggcgcggttagatgtcaatctgtctgcaggttggaagaatccaactctgtcgtactgacaag	50359
Qy	1088	aaaaactctcaagaaacgtctctgtcaaaaactgtgtgtgtgagtagcgtttatccatgtttgag	1147
Db	50360	aaaaactctcaagaaagcgtgtgtggaaaaaactatccaggtgcgtgttgcacatctcaaggtgtgtgtg	50419
Qy	1148	cgtgcactgaagaaactgaataatgaagaaggaagaagatcgtgtgaagaatgctccaatgtgcaa	1207
Db	50420	cgtgcactgaagaaactgaatgtgaagaaggaaaaaagaccgtgtgtgaagcatgtcacttcaagcta	50479
Qy	1208	caagagctgcggtctgaaagaaggtatgttccctgtgtgtgtgatactgttctgtccgtcca	1267


```

Db 1213 aaaggttcgaacatgctcctcgaagctgtcgaaaggtatgaatttcaaccgcgataacta 1272
Qy 601 tctcatactcttgytaactaatcccgagaataatggttctgtaacttgtaacccttatac 660
Db 1273 tccagctactcttctcaacaatccgaaacacagaatgttctttagaagaagctctcgtg 1332
Qy 661 ctcttgtaatgagaataatctactagcaatgaagaacatgtaaccatctttagaacaagt 720
Db 1333 ctatctctgtataaaaaaatttccggaatcaagaatttcttaccaggttttacaacaa 1392
Qy 721 gctaaaglaaacgcgtccactcttattatgtctgaagaagctagaaggtgaagcacttca 780
Db 1393 gcagaatcagaagcgtccctacttactatctgtcgaagatcagaaggaagcactttagct 1452
Qy 781 acactttagtcaatgaagctcgttgaagcactccaagtttagccgtaaaagctcctgtg 840
Db 1453 actttagtagtaaacacagactcgtcgtgattcagagtggtgtgcaataaaagctcclgga 1512
Qy 841 ttgtgtgaagcgcgttaagcctatgcttgaagatagttctacttactctggaaggaagca 900
Db 1513 ttgtgtgaatcaacaacaaagctatgttagaagacatcgcctatttcaactgtgtgtaactc 1572
Qy 901 atatttgaagaatcgttggatataaaagcttgaataatgtaagcttgcctctttagaagaag 960
Db 1573 atcagcgaagagcgttgcgtatgaagcttgaagaacacaaactctagctatgtttagaagaag 1632
Qy 961 aaagctgtgattatgacaaagaataactactactatcgttgaatgtgtcgtggaataatcaga 1020
Db 1633 aaaaaagctatcgttcttcaaaagaatatacaacaattgttgaagcttcttgcgaagaaga 1692
Qy 1021 gataatgaagctcgtagcttaacaacatcgcgtcgaacaattgaagaataaagctcagatatt 1080
Db 1693 gatcttgaatcgcgtcgtcgaagaatatacaaaaacaaatcgaagcagtaactctgatactac 1752
Qy 1081 gatcgtgaaaaaacttcaagaagcgtcttgcacaacacttgttgaagtagctgtatccat 1140
Db 1753 gacaaagaataaactccacagaagcgtttagctaaacttccggaagcgtgagctgtatccatcgt 1812
Qy 1141 gttggaagctcgtactgaactgaatgaatgaagaagaagatcgttgaagaatgtcctca 1200
Db 1813 gtaggaagctcgtactgaagaatgaatgaagaataaagaacagagtagatgtatgctcag 1872
Qy 1201 aatgcaacaagagcgtcgttgaagaaggtatgttcccttgggtgtgtactgttctgtgc 1260
Db 1873 catgcaactcttgcgtgaagttgaagaaggtatcttaccctgtgcgtgtgtacagctttagt 1932
Qy 1261 cgttccaattaaagtccttgaatgaat-----taaacctgtcgtatgaatgaatgtcgt 1314
Db 1933 cgtcgtacccctactttagaagcttcttacccttacttcaaatgaagaatgtgcaaatc 1992
Qy 1315 ggaactataatcatcgtcgttcttcttgaagaagcctttagcgtcaaatgtcgtcgaatgtc 1374
Db 1993 ggaagcagctatgttcttcaagaacatataccgtctcatttaagaagcagcaagct 2052
Qy 1375 ggcataatgaaggtctatgttctgaagaagaagctcgttgaaccaaaagtggttttggatt 1434
Db 2053 gataaagaagaagcgtatcatctgttcaacaagatgcttcttgcctccctttagcgaaggtat 2112
Qy 1435 aatgctgcatcagaagaataatgaagccttataaagctgtgtgtatgttcaatccataa 1494
Db 2113 gatgtcttaacgagcttcttaccacagcatatgttagagcaagaattctcgtcccaactaa 2172
Qy 1495 gtlacacgtatgtcatttacaataatgagcatcagtcagtccttacttcaactacaaga 1554
Db 2173 gtlacacgttgtgtcatttacaagaagcagcttctgttagctgtgtccttattataacaga 2232
Qy 1555 tgcgtctatgtcgtgaaaaaacagaaactaa 1584
Db 2233 gctttaatgtcgcgaatcttccgtgaagaaga 2262

```

RESULT 5
AA013137

```

ID AA013137 standard; DNA; 2223 BP.
XX
AC AA013137;
XX
DT 22-Oct-1991 (first entry)
XX
DE Hyp operon.
XX
KW Antibodies; heat shock; hypersensitive; allergen; HSP60; GroEL; ss.
XX
OS Chlamydia trachomatis serovar A.
XX
PH Key
FT CDS
FT 35..343
FT /*tag= a
FT /label= hypA
FT CDS
FT 361..2015
FT /*tag= b
FT /label= hypB
FT RBS
FT 23..26
FT /*tag= c
FT RBS
FT 35..38
FT /*tag= d
FT terminator
FT 2046..2072
FT /*tag= e
FT /note= "dyad symmetry; rho-independent"
FT
FT US7531317-A.
PN
XX
PD 09-JUL-1991.
XX
PF 31-MAY-1990; 90US-0143560.
XX
PR 31-MAY-1990; 90US-0531317.
XX
PA (USSH ) NAT INST OF HEALTH.
XX
XX WPI; 1991-245693/33.
XX DR P-PSDB; AAR13336, AAR13337.
XX
XX DNA encoding HypA and HypB Chlamydia proteins - used to develop
XX prods. for detection of and vaccines against Chlamydia infection.
XX
XX PS Disclosure; Fig 7; 51pp; English.
XX
XX The sequence was obtd. from clone pTA571 isolated from a DNA library
XX prepd. from C. trachomatis genomic DNA. The two ORFs encode HypA
XX and HypB hypersensitivity proteins. The DNA can be used to produce
XX recombinant proteins or to design probes for the detection of
XX Chlamydial infection. The Hyp proteins can be used to raise
XX antibodies and in vaccines. The Hyp B protein elicits a cell-
XX mediated immune response so can be used as a skin test antigen.
XX See also AA013136.
XX
SO Sequence 2223 BP; 745 A; 408 C; 489 G; 581 T; 0 other;

```

Query Match 43.1%; Score 710.4; DB 12; Length 2223;
Best local similarity 65.9%; Pred. No. 2.4e-153;
Matches 1048; Conservative 0; Mismatches 536; Indels 6; Gaps 1;

```

Qy 1 atggtcttaagaataccttctttagatgctaaagcccggtgaataacttcaagaggtgta 60
Db 381 atgttcgtctaaataaataatatacaacagaagaagccagaagaataatcaaaagaggt 440
Qy 61 gataaactgcaaatgtctttaaagtaaacacttgagacccaagaagccgtaatgctgtatt 120
Db 441 aagaacttgagcgaagcgtttaaagctactcaggcctaaaggaagacatgtgttcata 500
Qy 121 gaaaagctcttgggtcccgattatcaaaaagagtgctatcgttgcgaagaagaatt 180
Db 501 gataaagcttcgagatcccccataactaataagaatgcttaccgttgcgaagaagaatt 560

```

OY	181	gaacttgaaggttaagtttggaaaatattggggccccaatgtttgaaagatgtatcccaa	240
Dh	561	gagcttgcgcacaacaactgaaaataatlygggctccaaatggtcaaaagatctgccagaa	620
OY	241	actagcgaalattgcgttctgaatggaactacaacagcaacgctctgtccaaagcttat	300
Dh	621	actgcgtcaaaagcttggaaagccgaactacacacgcctaccgttctctgtcgaagctctat	680
OY	301	cgtgaaggtgaaaaacttctgaagcgtggtcgtaatccctactgaatgaatgaatgta	360
Dh	681	acagaagaatattacgcaatgtgaacagcttggcaaatccaaatgagctccaaacgagtatt	740
OY	361	gataaagctgttctgtctgttactaaagaaactaagcagcatatacaagcctactctgtac	420
Dh	741	gataaaagctgttaaggttctgttctgtatccaatcaaaaataatcagaagcctgtttacgat	800
~	801	caaaaaaataaagctcaagcttggaaaccttctgcacaactctgcatacaacaatgaat	860
OY	481	atcatacgtgaagctatgctgaagcttaaaagtggaaaaggggtgtattacaagtttgaagaact	540
Dh	861	ctgattctgctgaagcaatggagaaagtctgttlaaaaacggtctatcaatcgttgaagaaga	920
OY	541	aaaggtccttgaactacatacatagaatgtgttgaagaaatgaagtttgaccgtgtgcactc	600
Dh	921	aaaagattcttgaacccgttctggatctgttcttgaagaaatgaatttcaatataaggttacttc	980
OY	601	ctccatcacttcttaactaatccctggaaaatagtttctggaacttgataaaccctataatc	660
Dh	981	ctctgcctactctgcgaacaacaatccagaactccaagaatgtgtatgtgaagacgcttctgt	1040
OY	661	ctttgttaattgaagaaaaagatttacttaacatgaaagacacatgcctcaactctgaacaagt	720
Dh	1041	ctaatctcagcataaagaataattctctggtatccaagaatllctctccattttacaacaagt	1100
OY	721	gctaaagtaaacccgtcccaactctctaatctgtctgaagacgtgaagaagtgaagcacttga	780
Dh	1101	gctgaatcccgccgctccctcttcttattatagaagaagaacatgtgaagcgaactttagct	1160
OY	781	acacttcttgatcaataaagcttccgttggagaccccaagttgttagccgttaaaagctccgt	840
Dh	1161	acttctgctgtgaacagaatctcgttggatgattccgggttctgcacgttaaaagctccagc	1220
~	841	cttctgtaaaagccgctgaagatgatctgtcttgaagatattgtatcttcaatccctgaaggagaa	900
~	1221	cttggagaaatgaagaaaagcgtatgttggaaagacatcgctatcttaacccgggcgtccaaatc	1280
OY	901	atacttgaagacgtcggtatbaaagcgttgaaaaatgtaaagctgtctctcttgaagacgct	960
Dh	1281	attagcgaagaagtttggcagatgaataattgaaaaacgcttaacttgaatgttattgttaaagct	1340
OY	961	aaacggtgtatttcttgaacaaaagaataacttactatcgttgaatgtgtgcgtgaaatcatgaa	1020
Dh	1341	aaaaaagtattcgttcttctaaagaagacacgacatcgttccgaagaaatgggtgtgaaaaagaa	1400
OY	1021	gataatgaagctcgaatgaatacaaatctgtgcacaaatitgaagaaacaagcttcgaatat	1080
Dh	1401	gctttgaagacgtctgttgcgaagaatcacaataaaacaatatgaagaacagcttcttcgattac	1460
OY	1081	gattcgtgaaaaactcccaagaacgctctctgcacaactgttctgtgtgaatgaatgcgttatccat	1140
Dh	1461	gataaagaataaactcccaagaagcgtcttctgttaagcctctctgtgtggagtagcagatcatgc	1520
OY	1141	gttggagctgtctacttgaacttgaatgtgaagaagaagaagatcgttgaagaagatgtccta	1200
Dh	1521	gttggagctgtcaacaagaattgtgaatgaagaagaagaagaagaatcgttgaatgaatgtccta	1580
OY	1201	aaatgcaacaagaactcctcgttgaagaagaattatgttccctcgtgtgtgtgtgtgtgtgtgt	1260
Dh	1581	catgtctacaaatcgtcgtcgttgaagaagaagaatllctctcgtgtgtgtgtgtgtgtgtgt	1640
OY	1261	cgtctcatttaaaagctctgatgatgaattaaacgt-----ctgatgatgatgaactgtct	1314

Accession	Sequence	Position
Db	1611 cgtctgataccctactctcttgaagcctcttgcacatgttactaaagaaatatagcaaat	1700
Oy	1315 ggaactaaataacacacgcgtctctctctgaagagcctttagcgaatatgtcgcgaatgtc	1374
Db	1701 ggaagctgcgcatgttttgaagaagctcttccgcctcttgaacaatatgtctgaacgca	1760
Oy	1375 ggcataaaggtctcatctgtttctgaagaaagttctgtaaccaaagatggtttgattc	1434
Db	1761 ggaataaaggtgtgatacctctccaaacgaatatactccgctctgcgaacgaagatat	1820
Oy	1435 aatgctgatacagaagaataatgaagacccataaagctgtgtcgttaagatccataaaa	1494
Db	1821 gatgcattgcgtatgtgataccaagataatgcttgaagccggtatcttagatccctgtaa	1880
Oy	1495 gtacacgataatgcataccaataatgaagcatcagtaacccctactactaactaactacaga	1554
Db	1881 gtaacccgctctgtttttagaagaagcgcgctcccgtagctgattactcttgcacacgaa	1940
Oy	1555 tgcgataatgctgaanaaacgaagactaaa	1584
Db	1941 gctctcaatgcagaagatccagaagaaa	1970

RESULT 6
ID AAX86155 standard; DNA; 1661 BP.
AC AAX86155;
XX
DT 22-SEP-1999 (first entry)
DE DNA encoding a Streptococcus pyogenes heat shock protein (Hsp)60-2.
RW Heat shock protein; Hsp60-2; immune response; immunological carrier;
KW cancer control; tumour; sarcoma; cancer; gene therapy; ss.
XX
OS Streptococcus pyogenes.
PN M09935270-A1.
PD 15-JUL-1999.
XX
PF 29-DEC-1998; 98MO-CA01203.
PR 31-DEC-1997; 97US-0001737.
XX
PA (STRE-) STRESSEN BIOTECHNOLOGIES CORP.
XX
PI Mizzen L, Wisniewski J;
DR WPI: 1999-430397/36.
DR P-PSDB; AAU23904.
XX
PT New nucleic acid encoding heat shock protein-60 from Streptococcus,
PT useful in vaccines, as carriers for other immunogens, as anticancer
PT agents and for diagnosis
XX
PS Claim 3; Fig 4A-B; 176pp; English.

The present sequence encodes a heat shock protein, designated Hsp60-2. The protein, its fragments, variants and fusion proteins, are used to elicit or enhance an immune response against Streptococcus, and to elicit a similar response to a target antigen fused to the protein. Unlike other immunological carriers, Hsp60 proteins are not immunosuppressive so provide an increased response to any conjugated or fused antigen. Also, where used for cancer control, they lack the side effects associated with endotoxins. They can also be used to detect specific antibodies and in treatment or prevention of tumours (e.g. sarcoma or cancers of breast, ovary, prostate, lung, pancreas or liver). The Hsp60 polynucleotide is used for recombinant production of the protein, as a source of primers and probes for detecting streptococcal standard hybridization/amplification assays, and

XX WPI: 1999-045171/04.
 DR New isolated *Enterococcus faecalis* polynucleotides and polypeptides
 XX - used to develop products for the detection of *Enterococcus* and for
 PT use in vaccines for prevention or attenuation of *Enterococcus*
 PT infection.
 PS Claim 1: Page 414-416; 2084pp; English.
 XX
 CC A computer readable medium has been developed which has recorded on it
 CC 982 nucleotide sequences isolated from the *Enterococcus faecalis* genome.
 CC AAX12938 to AAX13919 represent these nucleotide sequences which are
 CC primary nucleotide sequences, also known as contigs. The computer-based
 CC system can identify fragments of the *Enterococcus faecalis* genome with
 CC commercial importance. The products can be used to detect the presence
 of *Enterococcus faecalis* in samples. They can also be used for
 ~~~ diagnosing *Enterococcal* infection in an animal and monitoring  
 ~~~ progression of disease, and for identifying agents which can be used to  
 CC modulate the growth or pathogenicity of *Enterococcus faecalis*, or
 CC another related organism, in vivo or in vitro. In particular the
 CC polypeptides encoded by the *Enterococcus faecalis* nucleotide sequences
 CC can be used in vaccines to prevent or attenuate an *Enterococcal*
 CC infection.
 SQ Sequence 3625 BP; 1335 A; 542 C; 750 G; 993 T; 5 other;

Query Match 42.4%; Score 698.4; DB 20; Length 3625;
 Best Local Similarity 66.1%; Pred. No. 1.5e-150;
 Matches 1040; Conservative 0; Mismatches 527; Indels 6; Gaps 2;

QY 8 ctaagaacactctttgtgtaagccgtgaacaaattccacaggtgtgataaac 67
 DB 368 caaagaagtttaatttgcagaagatgcagtcgcagcaatgctacgcggagatagatgtat 427
 QY 68 tgcgaatgtctgtaagaacacttgcacaaagccgttaagtctgtatctgaagaat 127
 DB 428 tagcagatcagtgaaagcagacattagccctaaagtgctgaacgtgttcttgaaaat 487
 QY 128 ctttgcgtcccgatattacaagaatggtgtatctgtctcaaaagaattgaaactg 187
 DB 488 ctttgcgtcccgatattacaagaatggtgtatctgtctcaaaagaattgaaactg 547
 DB 188 aagaatagttgaaataatggtgcgaacatggttaagaagaatgctcccaaaactagag 247
 QY 548 aagaatcatttgaataacatggtgcgaacatggttaagaagaatgctcccaaaactagag 607
 QY 248 atattgctgtgtaagaaacacagacagacagctcttgcacaagctatttactgtaag 307
 DB 608 atattgctgtgtaagaaacacagacagacagctcttgcacaagctatttactgtaag 667
 QY 308 ggtgtaaaactgtcagcagctgtctgtatctcctaagccctaaagctgtgacagataag 367
 DB 668 gctlaaaaaacgtaactgtgtagcacaaccatctaggtatctgcctgggattgaaatag 727
 QY 368 cgtgtgtgtcttactaaagaactcagacatgacatacaagctactgtgaccaaagaag 427
 DB 728 caacaaaaacagcagtagaagaattacacaatattccactgtagtgtatcaaaaagaag 787
 QY 428 aatagctcaagttgaaacacattctgcacaactctgatacaagaatagtaaatatcag 487
 DB 788 cgatgtcacaaagtcgtgtctgtctcactcaggtctcga---aaaagtcgacaaatcag 844
 QY 488 ctgaagctcagctaaagtctgaaagaagaggtgtatcaccagcttgagaagctaaagctc 547
 DB 845 ccgaatgcacaaagttggtlaacgcagcgccglaattacacatgaaagaatcaaaagga 904
 QY 548 ttgaatacattatagatgtgtgtgtaagaatgaagttaacgtgtgctccctccat 607
 DB 905 ttgaaacagaatcagatgtgtgtgtaagaatgaagttaacgtgtgctccctccat 964
 QY 608 acttgaactaattccctgagaagaatggttctgtaacttgataacctatataccttctgta 667

DB 965 acatgttactgcacaacgataaaatggaagccgttctttagaaatccatatactatca 1024
 QY 668 atgagaagaagttacttagcagaagacatgctcccaactcttagaacaagtgtcgaag 727
 DB 1025 ccgacaacaaaatctcaaatatcagaatattcacttacttactttagaacaatcttacaac 1084
 QY 728 taaccgtccactccttacttacttacttacttacttacttacttacttacttacttactt 787
 DB 1085 aagccgtccacttacttacttacttacttacttacttacttacttacttacttacttactt 1144
 QY 788 tagtcaataagctccgtgtagcactcccaagcttgcagcgaataaagctccgtgttgg 847
 DB 1145 tagtcaataaagctccgtgtagcactcccaagcttgcagcgaataaagctccgtgttgg 1204
 QY 848 aagcgcgttaagctatgtctgtaagaatattgtcactcctcctgtagagaagaatattg 907
 DB 1205 accgcgcgaacgtagtctgtaagaatattgtcactcctcctgtagagaagaatattg 1264
 QY 908 aagatcgtgtgtaagaatattgtcactcctcctgtagagaagaatattg 967
 DB 1265 aagcgtttaggttagaagtaaaagacacacactatgaaacttgaagaatgcttagcaag 1324
 QY 968 tagtattgacaaagaataactatcctatcgttgcagtgctgtagaagaatcagaatatta 1027
 DB 1325 tagtgcgacaaagaataactatcctatcgttgcagtgctgtagaagaatcagaatatta 1384
 QY 1028 aagcgtttaggttagaagtaaaagacacacactatgaaacttgaagaatgcttagcaag 1087
 DB 1385 atgcgcgcgttacttacttacttacttacttacttacttacttacttacttacttactt 1444
 QY 1088 aaaaacttgaagaacgcttgcgaacactgtgtgtagaagtagcttactcctgtagag 1147
 DB 1445 aaaaacttgaagaacgcttgcgaacactgtgtgtagaagtagcttactcctgtagag 1504
 QY 1148 ctgtactgtaaaactgaataagaagaagaagatcgttgcagtagaagaatgcttactaa 1207
 DB 1505 ctgtactgtaaaactgaataagaagaagaagatcgttgcagtagaagaatgcttactaa 1564
 QY 1208 caagagctgcgttgcagaagaatgttccctgcgtgtgtgtgtgtgtgtgtgtgtgtgt 1267
 DB 1565 caagagctgcgttgcagaagaatgttccctgcgtgtgtgtgtgtgtgtgtgtgtgtgt 1624
 QY 1268 ttaagctcctgtagaataactgcgtgtagaagaatgttccctgcgtgtgtgtgtgtgtgt 1327
 DB 1625 tctgtcgaagctgcgtgcgcagaa---gctgaagcgatgtgtgcacacagggatcaagatg 1681
 QY 1328 tccgtcgttctctgcagaagccttactgaactgaactgtgtgcagaatgctgtgtaaggt 1387
 DB 1682 tccgtcgttctctgcagaagccttactgaactgaactgtgtgcagaatgctgtgtaaggt 1741
 QY 1388 ctatgtgtgtaaaaaagcttgcagcaaaagaagtggttggatttaagctgtcag 1447
 DB 1742 cagtgattgtgcacaaactaaaabtgtaacttagtaactgcgaatgcagtaag 1801
 QY 1448 gagaatgaagacacttactaaagctgtgtatcgtatgcataaaagaatcagctatg 1507
 DB 1802 gtagaagtgtaaacatggttgcagccggtatggttgacccaacaaagaatcgttctg 1861
 DB 1508 catlcaaaaatgcagcatcagtagcctccttacttacttacttacttacttacttactt 1567
 DB 1862 ccttacaataatgcagcttctgtcagcttacttacttacttacttacttacttacttactt 1921
 QY 1568 aaaaacccaagc 1580
 DB 1922 acaaacccaagc 1934

RESULT 8
 AAX86153
 ID AAX86153 standard; DNA: 1654 BP.
 XX
 AC AAX86153;

XX 22-SEP-1999 (first entry).
XX DNA encoding a Streptococcus pneumoniae heat shock protein (Hsp)60-2.
DE
XX
XX Heat shock protein: Hsp60-2; immune response; immunological carrier;
KW cancer control; tumour; sarcoma; cancer; gene therapy; ss.
XX
OS Streptococcus pneumoniae.
XX
XX MO9935270-A1.
XX
XX 15-JUL-1999.
XX
XX 29-DEC-1998: 98MO-CA01203.
XX
XX 31-DEC-1997: 97US-0001737.
XX
XX (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
XX Mizzen L, Wisniewski J;
XX
XX WPI: 1999-430397/36.
XX
XX P-PSDB: AAY23902.

PT New nucleic acid encoding heat shock protein-60 from Streptococcus,
PT useful in vaccines, as carriers for other Immunogens, as anticancer
PT agents and for diagnosis
PS
PS Claim 3: Fig 2A-B: 176pp; English.

XX The present sequence encodes a heat shock protein, designated Hsp60-2.
CC The protein, its fragments, variants and fusion proteins, are
CC used to elicit or enhance an immune response against Streptococcus,
CC and to elicit a similar response to a target antigen fused to the
CC protein. Unlike other immunological carriers, Hsp60 proteins are not
CC immunosuppressive so provide an increased response to any conjugated or
CC fused antigen. Also, where used for cancer control, they lack the side
CC effects associated with endotoxins. They can also be used to detect
CC specific antibodies and in treatment or prevention of tumors
CC (e.g. sarcoma or cancers of breast, ovary, prostate, lung, pancreas or
CC liver). The Hsp60 polynucleotide is used for recombinant production
CC of the protein, as a source of primers and probes for detecting
CC streptococci in standard hybridization/amplification assays, and
CC therapeutically in gene therapy vectors.

XX Sequence 1654 BP: 500 A; 326 C; 401 G; 427 T; 0 other;

Query Match 41.9%; Score 689.8; DB 20; Length 1654;
Best Local Similarity 65.8%; Pred. NO. 1.1e-148;
Matches 1035; Conservative 0; Mismatches 532; Indels 6; Gaps 2;

Y 8 cttaagaaatccttttgaagccggtgaagaaacttcaagagtgatgaataac 67
DB 19 caaagaataataatttcacatgagatgccggttcagcatatggtcgtgcatcc 78
QY 68 ttgcaatgctgttaagtaacacttgaacgttaagggcgttaatgctgttaagaa 127
DB 79 ttgcaatgctgttaagtaacacttgaacgttaagggcgttaatgctgttaagaa 138
QY 128 ctttgatccccatcttcaagaagtgtatcgttcgtaaaagaataatgaactg 187
DB 139 catcggttcaccccttgataccagagtgatgactatgacaaagaatltgaattag 198
QY 188 aagataaatttgaataatggtcgtcaaatgttcaagaagtaagctcccaactagcg 247
DB 199 aagacatttgaataatcgtgtgccaattgtatcagaagtagcttcaaaacaaatg 258
QY 248 atatgctggtgagatgaacacacagcaagcttcgtcacaaatattatgagag 307
DB 259 atatgctggtgagatgaacacacagcaagcttcgtcacaaatattatgagag 318

QY 308 gtgaataacttgaagcagctgtcgttaactcctatgagccattaaagtgcatagaataag 367
DB 319 gaatcaaaaacgttccacagagtgcaaatccaatcgtatctcgtcgtgagattgaacag 378
QY 368 ctgttctgtctgttactaaagaactaaagcgaatcattcaaaagcctactcgttaccaaaag 427
DB 379 cagttgcgacagcagttgaagcttctgtaaaacaacagctatccctctgttcataaagaag 438
QY 428 aaatagctcaagttggaacatttctgcaaacctcgtatacaacaataggtatataatag 487
DB 439 ctatgcctcaagttgcagcgtatcttctcgttga--aaagttgtgatactct 495
QY 488 ctgaagctatggtgaaggttggaagagtggtatatacagcttggaagctaaaggtc 547
DB 496 ctgaagcaatggaataaagtgtgcaaaagcgtgtcatcaacatgaagatcaagtgtta 555
QY 548 ttgaactacattagatgtgttggaaggaatgaagttgacagcttgacactcctccat 607
DB 556 tgaacaacagagcttgaagctcgtagaaggaatgacagcttgacagcttgacactcctcaagt 615
QY 608 acttgttaactaactcctgaagaataatgttctgtaactcgtataacccctatattcttga 667
DB 616 acatgtgtacagatagcgaataaagtgtggtgaccttgaaatccgtatatttgatla 675
QY 668 atgagaataagattactagcatgaaagacatgctacacatcttgaacaagtgtgtaag 727
DB 676 cagacaataaataatccaatccaaagaaacttgcacatttgaagaacatctccaa 735
QY 728 taaacgctcacctcccttattcttgcgaagcgtgaagagtgaaacatttgaacacttg 787
DB 736 gcaatgcacactctgttatttctgtgagatgtgtgagtgagcgtcttccaaactctg 795
QY 788 tagtcaataagctcgtgtgagacatccaaagtgtgagccgtgaagaagctcgtgttgtg 847
DB 796 ttgtgaataaagattgttgaacacttcaacgtatagcagtaagcagctgtgttgtg 855
QY 848 aacgcgtgaagctatgcttgaagaatgtatgtcattccttcaggaaggaacaacttg 907
DB 856 accgtgcgaagcgtcgttgaagatatacgcacattcaacaggcgaacagttatcaag 915
QY 908 aagatcgtgtataaagcttgaataatgttaagctgttcttcttgaagaacagctaaacgt 967
DB 916 aagacctgtccttgaagtgtaagaagtcgcaaatgaagctccttgcataagcagagag 975
QY 968 tagtattgtcaaaaataactactatcgttgaatgtgtcgtggaataatcagaagaatata 1027
DB 976 tgaccgtgacaataagatagccagttatgttagaaggtgcaggaatcctgaagcgtatt 1035
QY 1028 aagctcgagttcaaaaacttgcgtcaaatgtgaagaacagctcagattatgtgtg 1087
DB 1036 ctcaacgtgtgtcgtgttcaagctcacaatccgaacactcaactcgtaatltgacgtg 1095
QY 1088 aaaaacttcaagaaagcttgcgaactgttgtgtgagtagctgttatacattgttgaag 1147
DB 1096 aaaaattgcaagaagcgttgcgaactgttgcaggtgtgtgtgagtagctgttataagtcgag 1155
QY 1148 ctgtcactgaaacctgaataatgaagaagaagatcgtgttagaagaatgtccttaataatgca 1207
DB 1156 ccgcacaactgaaactgaatgaagaagaagaatcgcgaatgaaatgtccctcaacgtta 1215
QY 1208 caaagctgcgtgtgaagaagaatgtatgtccctgcgtgtgtgtatcgttctcgtcctca 1267
DB 1216 ctgtgtcagcgtgttgaagaagaatgtatgttgcaggtgtgtgtgaagaagccttctgcaatgtga 1275
QY 1268 ttaaaagcttgaatgataltaaaactcgtgtgaatgaatcgttgcgaactaataatca 1327
DB 1276 ttccagctgtgtactccttgaatgtgacagagtagaagca--acagagatgaattctg 1332
QY 1328 tccgtgtccttcttgaagaagccttgaatgaatgtcgtcgaatgtcgtcgtatgagagtt 1387
DB 1333 tctcgtgtccttgaagaagccttgcgtcaaatgtcacaatgacagagatttgaagat 1392
QY 1388 ctattgtgtgaagaagctcgtgaaccaaagaatgttcttgatttaagtgtcgtacag 1447

| | | | |
|----------|---|--|------|
| Db | 1393 | ctaccgttatacgaatcgttllgaaaaaagcagagcttggtataggaatccaaagcagcaactg | 1452 |
| QY | 1448 | gagataatgaagacctataaagctgtgtcatatgattcctaataaaagtacagctatcg | 1507 |
| Db | 1453 | gcggagtggtgtaataatgatattgatacgaagatatacattgacatgaagagagtcgtccg | 1512 |
| QY | 1508 | catlacaataaalgacagcatcagtagcctcttacttacttactaacaagaalgcgtatctg | 1567 |
| Db | 1513 | ccttacaataatgcagctctcgtacgcagctgtattttgacaacaagaagcagtcgtgcca | 1572 |
| QY | 1568 | aaaaaacagaacc | 1580 |
| Db | 1573 | ataaacccagaacc | 1585 |
| UL* | 9 | | |
| 25036 | | | |
| 1U | AAE25036 | standard; DNA; 1926 BP. | |
| XX | AAE25036; | | |
| XX | 30-APR-2001 | (first entry) | |
| XX | | | |
| DE | | Nucleotide sequence of Hsp65-E7 fusion protein. | |
| XX | | | |
| KM | | Heat shock protein; HSP: Th1 response; Th1 cell; CD4+ T lymphocyte cell; | |
| KW | | Lymphocyte; Hsp65; Hsp40; Hsp10; Hsp60; Hsp71; microbial pathogen; | |
| KW | | E7 protein; ss. | |
| XX | | | |
| OS | | Synthetic. | |
| OS | | Streptococcus pneumoniae. | |
| OS | | Human papillomavirus. | |
| XX | | | |
| FT | Key | Location/Qualifiers | |
| FT | CDS | 1..1926 | |
| FT | | /*tag= a | |
| FT | | /product= "Hsp65-E7 fusion protein" | |
| XX | | | |
| PN | | WO200104344-A2. | |
| XX | | | |
| PD | | 18-JAN-2001. | |
| XX | | | |
| PF | | 10-JUL-2000; 2000MO-US18828. | |
| | | | |
| | | 08-JUL-1999; 99US-0143757. | |
| ..A | | | |
| PA | | (STRE-) STRESSGEN BIOTECHNOLOGIES CORP. | |
| XX | | | |
| PI | | Siegel M, Chu NR, Mizzen LA; | |
| XX | | | |
| DR | | WPI: 2001-138361/14. | |
| XX | | P-PSDB: AAB31619. | |
| PT | | | |
| PT | | Screening for compounds that stimulate Th1-like responses in CD4+ T | |
| PT | | lymphocyte cells | |
| XX | | | |
| PS | | Example 15; Fig 15A-B; 88pp; English. | |
| XX | | | |
| CC | | The present sequence encodes a fusion protein comprising a | |
| CC | | Streptococcus pneumoniae heat shock protein (Hsp) 65 fused to a HPV16 E7 | |
| CC | | protein. The fusion protein is used in the method of the invention. The | |
| CC | | specification describes a method of determining whether a compound | |
| CC | | stimulates a Th1-like response. Th1 cells are a subset of CD4+ | |
| CC | | T lymphocyte cells. The method comprises contacting naive lymphocytes | |
| CC | | in vitro with a fusion protein comprising at least a fragment of Hsp, | |
| CC | | and then detecting the Th1-like response exhibited by the cell sample. | |
| CC | | The proteins which may be used in the method of the invention are Hsp65, | |
| CC | | Hsp40, Hsp10, Hsp60, and Hsp71. The method may be used to identify | |
| CC | | compounds that stimulate Th1-like responses in response to microbial | |
| CC | | pathogens. | |
| XX | | | |
| Sequence | 1926 BP; 585 A; 382 C; 464 G; 495 T; 0 other; | | |

| | | | | | |
|-----------------------|--------------|--|----------------|-------------|---|
| Query Match | 41.9% | Score 669.8 | DB 22 | Length 1926 | |
| Best Local Similarity | 65.8% | Pred. No. 1.2e-148 | | | |
| Matches 1035 | Conservative | 0 | Mismatches 532 | Indels | 6 |
| Gaps | | | | | 2 |
| QY | 8 | ctaaagaaatccctttttagatgctaaagcccggtgaaaacattccacgaggtgataaac | 67 | | |
| | | | | | |
| DB | 5 | caaaagaaattaaatttcaatcagatgccggttcagctatggtccgtggtcgtgataacc | 64 | | |
| QY | 68 | tttgcacatgctgttaaagaaacacattggaactcgaacaaagccgttaagtctgtattgaaagt | 127 | | |
| | | | | | |
| DB | 65 | ttgcgcgatactgttaaagaaacttttgggcccaaaaggtgcgaatgtcgttcttgaaagt | 124 | | |
| QY | 128 | cttttgttccccattatcaaaaagatggtgatactgtgtgcaaaagaattgaaactgt | 187 | | |
| | | | | | |
| DB | 125 | catctggttcaacctttagattaccatgacgtgtgtgactatgtccaaaagaattgatatag | 184 | | |
| QY | 168 | aagataagtttgaataatagggcgttcacaaatgtgttcaagaagttagtctccaaaactaagg | 247 | | |
| | | | | | |
| DB | 185 | aagaccattttgaaataataggggtgccaaatvtgatacgaagtagtgcttaaaaacaaatg | 244 | | |
| QY | 248 | atatgctgtgtatgtgaacacaacagacacagctcccttcacagaactatttcgtgaag | 307 | | |
| | | | | | |
| DB | 245 | atatgcgcgggtgtatgtgaactacaactgcacacgttttggcccagaacatcgtccgtaag | 304 | | |
| QY | 308 | gtgtaaaacctgttagcagcgtcgtcgaatccatactgagccataaocgttgacatagataag | 367 | | |
| | | | | | |
| DB | 305 | gatacaaaaacgtcacacagaggtgtcacaatcccaatcgttatctcgtgtgatatgtgaaacg | 364 | | |
| QY | 368 | ctgtgtgtgtcttacttaagaacttaagcgacataccaagaagctactcgttgaccaaaag | 427 | | |
| | | | | | |
| DB | 365 | cgatgtgcgcgaacgttgaaagcttctgaaaaaacaacgttccctgttgcataaagaag | 424 | | |
| QY | 428 | aaatgctgaagtgtgaacacattcttcgcacaactctgatacaagaatgataatcatag | 487 | | |
| | | | | | |
| DB | 425 | ctatcgtccaaagtgtcagcgcgtatcttctcgttctga-----aaagtgtgtgataatct | 481 | | |
| QY | 488 | ctgaagctatgtgtctaaagtttgaaaaagaggtgtgtacacagtttgaggaagctaaaggtc | 547 | | |
| | | | | | |
| DB | 482 | ctgaagcaatgtgaaaaaagtttggcaaaagacggtgtcatccatcgaagggtcaggtgtga | 541 | | |
| QY | 548 | ttgaaactacatlagatgtgtgtgaaggaatgaagtttgcacgtgtgtaccctctccat | 607 | | |
| | | | | | |
| DB | 542 | tgtgaaacagagcttgaaagtcgtgaagaagatcagtttgcacgtgtgtacccttccacgt | 601 | | |
| QY | 608 | acttgttaactatccctcgnaaaaatggtgttggaaacttgataacccttatcccttga | 667 | | |
| | | | | | |
| DB | 602 | acatgtgtgcagatagcgaaaaaatggtgtgtgcgttgcacttgaaaatccgttcatattgtata | 661 | | |
| QY | 668 | atgagaaaaaaggttactaagcgtgaagacatgtctacaaactcttagaacaagaattgtcctaag | 727 | | |
| | | | | | |
| DB | 662 | cagacaagaabaatttcccaatataccaagaataatctgcgaacttttggaaagcatcttccaa | 721 | | |
| QY | 728 | taaacgctaccctcttattatgtctgaaagcgtlagaaggtgaaagcattgcacaacatgt | 787 | | |
| | | | | | |
| DB | 722 | gcaatcgtccacactctgtatattgtcgtgatagtgtgtagtgaaggtcttccacaactgtg | 781 | | |
| QY | 788 | tagtcaatbagctcgttggagcacttccaagttgttagcgttaaaagctctgtttgtttgt | 847 | | |
| | | | | | |
| DB | 782 | ttttgaaacaagattcgttggaaaccttcaacgtltagtcaacgaaaggaacctgtgtttgtgt | 841 | | |
| QY | 848 | aacgcgcgtaaagctatgctgtgaagatactgtctacccttctgtgaaggagaagcaataattg | 907 | | |
| | | | | | |
| DB | 842 | accgttcgcaaaagcactgtcttgaagatatatgccttcaacagcgcaaacgttatcaag | 901 | | |
| QY | 908 | aagatcgtgtgataaaagcttgaaaatgtlaagcttgtctcttltaggaacagctaaaagtg | 967 | | |
| | | | | | |
| DB | 902 | aagacctgtgtctgtgtgtgaagaatgtgacaaattgaagctctgtgttcaagcagcgagag | 961 | | |
| QY | 968 | tagtatactgacaagaataatctactatcgttgaatgtgtcgtgaaaatacgaagaatcata | 1027 | | |
| | | | | | |
| DB | 962 | tgaacgtgtgacaagaatgaacaggtatctgttgaagaaggttgaagaaatcttgcgaagcgaatt | 1021 | | |

| | | | |
|-----------|--|---|------|
| QY | 1028 | aagctcgaagcttaaaacaaattcgcgcacaaattgagaagaacagcgcagattatgctg | 1087 |
| Db | 1022 | ctcacctgcttcgctgtatccaagctccaacacgcgaactcaactcctaattgaccctg | 1081 |
| QY | 1088 | aaaaaacctcagaagaacgctctgcacaacactctgttgcgtgcgaatcgttaccatgctgag | 1147 |
| Db | 1082 | aaaaattgcagaagaacgcttggccaaatttcgaagtcggtgtacggttattaagtcgag | 1141 |
| QY | 1148 | ctctctactgaacatctgaaatgaaagagatgcgtgtacgaagaatgcctctaataatgaa | 1207 |
| Db | 1142 | ccgcgaactgaacatctgaaatgaaagaaatgaaactccgcacattgaagatccctccaacgcta | 1201 |
| QY | 1208 | caagagctgcgcttgcgaagaaggatattgtccctgcgtgcgtgtactgtcttctgcgtcca | 1267 |
| Db | 1202 | ctcgtgcagctctgttgcgaagaaggatattgtctgcaggtgcgtgcgaacagctcttgcgaatgta | 1261 |
| QY | 1368 | ctaaagctcctctgatagtatlaaactccctgcctgcatagcatagtaaacttgcgtactaata | 1377 |
| Db | 1262 | ctccagctgtgtgtactcctctggaattgcagaagatgaagca---acaggaactaattatg | 1318 |
| QY | 1328 | ctccgtctctctcttgcgaagccttaccgtccaatctgcgtgcgaatgcgtgcgtgaaggtc | 1387 |
| Db | 1319 | ctctccgctgcttgcgaagaacactgctgcgaactatgctacacatgcagatattgaagat | 1378 |
| QY | 1388 | ctatttgttgtgagaanaagcttcgtgaacccaagaatggttttgatttatgtctgcatcag | 1447 |
| Db | 1379 | ctctcgttactcgcgtcttcgtaaaaagctgcgaagctcgtgataagattcaacgcagcaactg | 1438 |
| QY | 1448 | gagaataatgagaacacttataaagctgcgtgcattgcatctgatacttcaaaaagcttacaagatg | 1507 |
| Db | 1439 | gcgagtggtgttaacatgatattgataccaagatgatactgataccagttaaagtgaagtcgtcag | 1498 |
| QY | 1508 | gattacaataatgcagcatcagtagctcctccttactcttcaactacagaatgcgtatgctg | 1567 |
| Db | 1499 | ccctacacaaatgcagatctgtgacagacttatttgcacaacagaagacagctgtaagcca | 1558 |
| QY | 1568 | aaaaaacagagac 1580 | |
| Db | 1559 | ataaacacagagac 1571 | |
| RESULT 10 | | | |
| AAV52210 | | | |
| ID | AAV52210 | standard: DNA, 5365 | BP. |
| XX | AAV52210: | | |
| AC | | | |
| XX | | | |
| DT | 23-OCT-1998 | (first entry) | |
| XX | | | |
| DE | Streptococcus pneumoniae genome fragment SEQ ID NO:77. | | |
| XX | Streptococcus pneumoniae: S. pneumoniae: genome: diagnosis: assay: | | |
| XX | computer readable medium: vaccine: pharmaceutical composition: ds. | | |
| OS | Streptococcus pneumoniae. | | |
| XX | | | |
| PN | NO9818931-A2. | | |
| XX | | | |
| PD | 07-MAY-1998. | | |
| XX | | | |
| PF | 30-OCT-1997: 97WO-US19588. | | |
| XX | | | |
| PR | 31-OCT-1996: 96US-0029960. | | |
| XX | | | |
| PA | (HUMA-) HUMAN GENOME SCI INC. | | |
| XX | | | |
| PI | Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M; | | |
| XX | Kunsch CA, Rosen CA; | | |
| XX | WPI: 1998-272225/24. | | |
| XX | | | |
| DT | Computer-readable medium with recorded Streptococcus pneumoniae | | |

| | |
|-----------------------|---|
| PT | polynucleotide sequences - useful in diagnostic kits and assays, and |
| Pr | pharmaceutical compositions and vaccines for Streptococcus |
| Pt | pneumoniae |
| XX | |
| XX | Claim 1; Page 628-631; 1409pp; English. |
| XX | |
| CC | The present invention describes a computer readable medium which has |
| CC | the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524) |
| CC | recorded on it, or a representative fragment or a sequence at least 95% |
| CC | identical to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in |
| CC | SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from |
| CC | Streptococcus pneumoniae. The present invention also describes an |
| CC | isolated nucleic acid molecule encoding a homologue of any of the |
| CC | fragments of the S. pneumoniae genome (SEQ ID NO:1 to 391) where the |
| CC | nucleic acid molecule is produced by a process comprising: (a) screening |
| CC | a genomic DNA library using as a probe a target sequence defined by any |
| CC | of the sequences in SEQ ID NO:1 to 391, identifying members of the |
| CC | library which contain sequences that hybridise to the target sequence and |
| CC | isolating the nucleic acid molecules from the members; or (b) isolating |
| CC | mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid |
| CC | molecules whose nucleotide sequence is homologous to amplification |
| CC | primers derived from the fragment of the S. pneumoniae genome to prime |
| CC | the amplification and isolating the amplified sequences. The computer |
| CC | readable medium can be used in a computer-based system for identifying |
| CC | fragments of the S. pneumoniae genome of commercial importance, or |
| CC | expression modulating fragments of the S. pneumoniae genome. Products |
| CC | from the present invention can be used in diagnosis kits and assays, and |
| CC | pharmaceutical compositions and vaccines for S. pneumoniae. |
| XX | |
| SQ | Sequence 5365 BP; 1654 A; 960 C; 1082 G; 1669 T; 0 other: |
| Query Match | 41.7%; Score 687.6; DB 19; Length 5365; |
| Best Local Similarity | 65.7%; Pred. NO.5,1e-148; |
| Matches 1034; | Conservative 0; Mismatches 534; Indels 6; Gaps 2. |
| OY | 7 tctaagaacatccttttgatgctaaagccgtgaaaaacttcaacgagtgatgataaa 66 |
| DB | 1 tttttttt ttt tttttttt ttt tttttttt |
| DB | 281 tcaaaagaacattcaatttcacatgagtcgccgttcacatgcttcgtgtgcgatac 340 |
| OY | 67 ctccaagtgcgttaagaatgaactcttgagccttaagagccgtatgctttttaaag 126 |
| DB | 341 ctggaagacactgtttaagaatgaactccttgaggacaaaagtcggaatgcttcgtgaaag 400 |
| OY | 127 tctttggttcccccagttatcaaaaagatggtgtatcgtttgtgcacaaagaattgaact 186 |
| DB | 401 tcatcgtgtccacccttgattacaatgacggtgtgacatcgtgccaaagaatcgatcg 460 |
| OY | 187 gaagataagtttgaataatgtaggctccaatggtgttaagaagaagtgtgtcccaaatgac 246 |
| DB | 461 gaagacccttttgaataatgtaggtgtcctaagttagtacaagaagtgtcttcaaaacat 520 |
| OY | 247 gataatctgtgtatgataactcaacaacagcaacagctctgtgcacagctattatctgtga 306 |
| DB | 521 gatacgcaggttgacggagacactaagctgacaacagctcttgaccacaagctatgctcgtaga 580 |
| OY | 307 ggtgtaaacacttgcagcgtcggtcgttaacctcattgycatcaaaacgtgacatagataaa 366 |
| DB | 581 ggaatcaaaaacgtcacacgaggtgcataatcaaatcgcgtattcgtctgtgattgaaca 640 |
| OY | 367 gctctgtgtgttactaacaagaactaagcgacattacaagctctactctgtgacgaataa 426 |
| DB | 641 gcaagtgcgcagcagcttgaaagcctttgaaaaaacaacgcatccctcgtttgccaataaaga 700 |
| OY | 427 gaatatgctcaagtgtgacaacttctgcacaactctgtatatacaaatagatatacata 486 |
| DB | 701 gctatcgcataagtgcagcgcgtatctctcgtctga---aaaagtgtgtagatacacc 757 |
| OY | 487 gctgaagctatgctcaaaaggtgaaaagaggtgtatcacaacgtgtgaggaactaaagt 546 |
| DB | 758 tctgaagcaatgaaaaaagcttgcgaagaagcgtgtcaccacacacgcgaagaagtcacagtgt 817 |
| OY | 547 ctgaatacatatgatagtgttgaagaagaagtttgcacgtgtgactctctcaca 606 |


```

DB 342835 acgctgcacaaagctgagcgaactacacacgctacatgcttcttctgtaagctatctat 342894
QY 301 cgtgaagctgtaaaactctgttagcagctgttcgttaactcttaagcctaataacgttgcata 360
DB 342895 acagaaagallacgcaaltgaacagctgagcaaatccaaatgaaacccaaagaagatatt 342954
QY 361 gataaagctgcttctgttcttactaaagaactaagcagatcaacaaagctactctgtac 420
DB 342955 gataaagctgcttgaagctgttctgttgaatcaaaaataatcagcagcttcttcagat 343014
QY 421 caaaaagaataagctcaagcttggaacattcttcgaacactctgaacactgaacaacaaagtaac 480
DB 343015 caaaaagaataatctcgaagcttggaacaaattcttcgaacactgaacaaacaaagtaac 343074
QY 481 atcatagctgaagctatgagcgaagcttggaagagctgttaacaaagctgaagagct 540
DB 343075 ctgattgctaagcaatgaggaagcttggtaaaaacgctctatcaactgttgaagaagca 343134
QY 541 aaagctctgaaacatcattagatgtgtgtgaaggaatgaagcttgaacgctgctactc 600
DB 343135 aaaggaattgaaacgcttctgtgtgtgtgtgaaggaatgaatctcaatagagcttactc 343194
601 tctccatctgttgaactaactctgagaaatggttctggaactgtatgaaccttatac 660
DB 343195 tctagctactctgcaacaataatccagaacactcaagaatgtgtatagaagaagcttctgt 343254
QY 661 cttgtgaatgagaaagaatctactagatgaagacatgacatcacaactctagaacaagt 720
DB 343255 ctactctagatagaagaatcttctggaatcaagaaltctctctctgttatacaacaagt 343314
QY 721 gctaaagtaaacgcttccactctctattatctgtgaagagctgaagaagctgaagcactgtca 780
DB 343315 gctgaatccggtccctctctctctctctctctctctctctctctctctctctctctct 343373
QY 781 acactgtgactcaataagctcgtgagcagctcgaagcttgtaacgttaaaagctctcgt 840
DB 343374 accttgtctgtgacagaagatctgtgagagatctcgggtctgcagcttaaaagctccagc 343433
QY 841 ttgtgtgaagcgcgtgaagctatgtctgaagatctgtctactctctctgaagagagaaga 900
DB 343434 ttgtgagatagaagaagaagctatgtctgaagacatcgtactctctctctctctctctct 343493
QY 901 atatttgaagatctgtgtataaagcttgaagaatgaagctgtctctctctctctctctctct 960
DB 343494 attagcgaagagctgtgagcagtaaatagaacaagcctaactgtatgtatgaagct 343553
QY 961 aaacgtgtgacttgaacaaagaataactactactactgtgtatgtgtctgtgaaatcaga 1020
DB 343554 aaataaagctatctctctctctctctctctctctctctctctctctctctctctctctct 343613
1021 gatataaagctgtgaatgaacaaatctgtgcacaaatgtgaagaacaaagctcagatatt 1080
DB 343614 gcttgaagcgtctgtgagaaagcattcaaaaacaaatgtgaagacagctctctctgtat 343673
QY 1081 gatcgtgaaaaaacttcaagaagctctgtcaaaactgtgtgtgtgtgtgtgtgtgtgtatccat 1140
DB 343674 gataaagaataaactccaagcgtctctgtcaagctctctgtgtgtgtgtgtgtgtgtgtgt 343733
QY 1141 gttgagcgtctactgaactgaatgaagaagaagaagaagcgtctgaagaagctgctctctct 1200
DB 343734 gttgagcgtctgaacagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 343793
QY 1201 aatgcacaagaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1260
DB 343794 catgcataaactcgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 343853
QY 1261 cgtctcaataaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1320
DB 343854 cgtctgtatccctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 343913
QY 1315 ggccttaataatcatccgctgtctctctgtgaagaagccttgaagctcaaatgtctgtgaagct 1374

```

```

DB 343914 ggaagctgcagattgtttgaaagctcttccgctctcttctgaaacaaatgtctgcaaacgca 343973
QY 1375 ggcctgaaggtctctatgtttgtagaaaaagctctggaaccaaagaatggtttgtatt 1434
DB 343974 ggaagaagaagctgtctatctctctcaacaagct--atgtctctctctggaacgaagatatt 344031
QY 1435 aatgcctcaggaagaataatgaaagccttattaaagcgtgtgtgtgtgtgtgtgtgtgtgtgt 1494
DB 344032 gatgcattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 344091
QY 1495 gtacacagctattgtctatacaaaatgacagatcagtagcctcctactctcaactacaga 1554
DB 344092 gtaaccgcttctgtctttagaagaagcgtctcgtctgtgtgtgtgtgtgtgtgtgtgtgt 344151
QY 1555 tgcgtctatgtctgtaaaaaccagaacctaa 1584
DB 344152 gctctcatgtcagagaatctccagaagaanaa 344181

RESULT 12
AAC90805
ID AAC90805 standard; DNA; 1647 BP.
XX
AC AAC90805;
XX
DT 16-MAR-2001 (first entry)
XX
DE Escherichia coli GroEL nucleotide sequence SEQ ID NO:2.
XX
KW Escherichia coli; GroEL; mutant; mutetin; improved stability;
KM chaperone; nontropic; neuroprotective; Alzheimer's disease;
KW prion disease; Creutzfeldt-Jacob's disease; CJD; ds.
XX
OS Escherichia coli.
XX
PN WO200072463-A1.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000MO-GB02019.
XX
PR 27-MAY-1999; 99GB-0012445.
XX
PA (MED-) MEDICAL RES COUNCIL.
XX
PI Buckle AM, Fersht A;
XX
DR MPI: 2001-061543/07.
XX
DR P-PSDB; AAB50536.
XX
PT Novel GroEL chaperone polypeptide, its homologue or fragment comprising
PT modifications in specified amino acid residues of GroEL protein
PT sequence, used to treat diseases associated with protein/polypeptide
PT structure
XX
PS Disclosure: Page 47-49; 58pp; English.
XX
XX
CC The present invention describes a GroEL chaperone polypeptide (CP) (I),
CC its homologue, or fragment having protein refolding activity, and
CC comprising one or more amino acid modifications at any one of amino
CC residues 207, 212, 217, 223, 233, 267, 271, 294, 305, 308 and 326 of the
CC fully defined GroEL protein sequence given in AAB50536, or their
CC equivalent positions in other homologous CPs. (I) can have nontropic and
CC neuroprotective activity. (I) can be used in therapy (i.e. gene
CC therapy). (I) is used in the manufacture of medicament for treating a
CC disease associated with protein/polypeptide structure. (I) immobilised to
CC a solid phase, on contact with a molecule/reconditions it, in which the
CC molecule is subjected to inactivation or denaturation prior to contact
CC with (I). The solid phase is a chromatographic matrix and the contact of
CC the molecule and (I) is carried out by applying the molecule to the top
CC of a bed of the matrix packed in a column and then eluting the molecule
CC through the column. (I) is useful for altering the structure of a
CC molecule by folding, unfolding or refolding. Preferably, the

```

CC stoichiometry between the chaperone (1) and the molecule being altered
 CC is about 1:1. (1) is also useful for purifying or increasing the yield,
 CC specific activity and/or quality of biological molecules. (1) can also
 CC be used for treating a animal or human patient suffering from a disease
 CC associated with protein or polypeptide structure. The polypeptides when
 CC administered either directly or via nucleic acid constructs may be
 CC useful for treating Alzheimer's disease and prion diseases including
 CC Creutzfeldt-Jacob's disease. The mutant or modified polypeptides have
 CC increased stability while retaining full chaperone activity. The present
 CC sequence encodes the Escherichia coli GroEL chaperone, which can be
 CC modified for use in the present invention.

XX Sequence 1647 BP; 437 A; 394 C; 477 G; 339 T; 0 other;

Query Match 41.68; Score 684.4; DB 22; Length 1647;
 %est Local Similarity 63.58; Pred. No. 1.9e-147;
 matches 1045; Conservative 0; Mismatches 601; Indels 0; Gaps 0;

OY 1 atggtcttaagaaactcttttgatgtctaaagccgtgaaaaactctcagagtgta 60
 DB 1 atggaagctaaagaagctaaatctcgtaacgagctcgtgtaaaatctgcgcgagct 60
 OY 61 gataaactgtcaaatgtctgttaagaactgtgaccctaaagccgtgacgtctat 120
 DB 61 aacgtacgtcagatgcagtgaaagttaacctcgtcccaaaagccgtgacgtatctgt 120
 OY 121 gaaaagctcttctgctcccaagtatacaaaagatggtatcgtgtgcaaaagaat 180
 DB 121 gataaatcttcgtgctcccaagctacccaagatggtgtcttcgtgctgctgtaaatc 180
 OY 181 gaactgtgaagtaagttgtgaataatagggcgctcaaatgtgttaagaagctcccaaa 240
 DB 181 gaactgtgaagcaagctgtgaataatagggcgctgagatggtgaagaagcttgcctctaa 240
 OY 241 actagcgatactgtcgtgtgatactgaacacagcaacagctcctgtcgaagctat 300
 DB 241 gcaaacgagcgtcgtcagagcagtgatacccaactgtcaacgtatcgtccagctatc 300
 OY 301 cgtgaaggtgtaaaactgtgacgctgtcgtatccttaagccatataacgtgacata 360
 DB 301 actgaaggtctgaagcgtgtcgtcgtgacgtgaacccgagtgacgtgaacgtgtgatac 360
 OY 361 gataaagctgtgtgctgtacttaagaactaaagcagatctataacagcctactcgtgac 420
 DB 361 gataaagcgtgtacccgctgagctgtgaagaactgaagcgtcgtcgtatcctacgtcgtac 420
 OY 421 caaaaagaatagctcaagttgtgaacacattctgtcaaaactgtatatacaatagta 480
 DB 421 tctaaagcgtatgtccaggtgtgtacatctccgttaactccgacgaacgttagtaaa 480
 OY 481 atcatagctgaagctatggtcaaaagtgtgaaagaggtgtatatacagttgaggaagc 540
 DB 481 ctgtacgtcgtgagcgtatgacaaagctgttaagaagcgttatcaacgttgaagagct 540
 OY 541 aaaggtctgtgaactacttagatgtgttgaagaagatgagttgacgcgtgactc 600
 DB 541 aacggtctgtgaacgacgtacgtggtgtgaaggtatgacgtcgtgacacgtg 600
 OY 601 tctccataactgttaactaatcctgtgagaaatgtgttgaactgtataaccctata 660
 DB 601 tctcttactatcaatacaacagcggaaactgtgcgagttagaactgtgaagccgtctac 660
 OY 661 ctltgtataggaagaatatacagatgaaagacatgcatcctaactcttaagaacagt 720
 DB 661 ctgtcgtgtgacaagaataatctccacaactccgcaaatgtcgtcgttctgtgaagcgt 720
 OY 721 gctaaagtaaacgcctcctcttatatgtgtgaagaagctgaagagctgacattca 780
 DB 721 gctaaagcaggaacacgcctcgtatcatcgtgtgaagatgttgaagaagcagcgtgaca 780
 OY 781 acaactgtatgaataaagctcgtgtgagcactccaagtgttagccgtataaagctcgt 840
 DB 781 atactgtatgaataaagctcgtgtgagcactccaagtgttagccgtataaagctcgt 840

DB 781 actcgtgttaaacacatcgctgacatcgtgaagaagctcgtcgtgttaagaacacgcgc 840
 OY 841 ttgttgtaagccgtaaaagctatgtcttgaagaatgtatctatcctacgtggagaaaca 900
 DB 841 ttgcgtgacgtcgtcaaaagctatgtctgacgatatctgaacccctgtacgtgcgtg 900
 OY 901 atattgtgaagctcgtgtatataaagcttgaatgtgaagctgtctctctttagaacaact 960
 DB 901 atctctgaagagatcgtatgtatgagctgtgaaagaacacccctgtgaagactgtgtcagct 960
 OY 961 aaacgtgtatgtatgtacaagaataatactactatcgtgtatgtgtgtgtgaataatcaga 1020
 DB 961 aaacgtgtgtatgtatcaacaagaacacacccactatcactcgtatgtgtgtgtgaagaagct 1020
 OY 1021 gatattaaagctcgtgataaaccaattcgtgacacaaattgtgaagaagaacacgtcagattat 1080
 DB 1021 gcaatccagggcgtgtgtcgtccagacccgtcagcagatgtgaagaacacactctgactac 1080
 OY 1081 gatcgtgaanaaactctcaagaacgctctctgcaaaactgtgtgtgtgtgtgtgtatcctat 1140
 DB 1081 gacgtgtgaanaaactctcaagaacgctcgtatgcgaacactgtgcagcagcgtgtgtcagttcctaa 1140
 OY 1141 gtltgagctgtactatgaactgaatgaagaagaagagatcgtgtgaagaatgtcctta 1200
 DB 1141 gt 1200
 OY 1201 aatgcaacaagaagcgtcgtgtgtgaagaagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1260
 DB 1201 cagcgtgacccgtcgtcgtgtgtgaagaagaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1260
 OY 1261 cgtccatataaagctcgtgtgtgtatataaactgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1320
 DB 1261 cgt 1320
 OY 1321 aatatcatcgtcgtctctctctgtgaagaagcctctatcgaacaaatgtctgtgaactat 1380
 DB 1321 aaagtgtacgtcgt 1380
 OY 1381 gaaggtcttatgtgtgtgtgaagaagaatcgtgtgaacaaagaatgtgtgtgtgtgtgtgt 1440
 DB 1381 gaacgtcgt 1440
 OY 1441 gcatcaggaagatatagaagccttataaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1500
 DB 1441 gcaacgtgaagatatagaagccttataaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1500
 OY 1501 cgtatgtatatacaaatgtcagatcagtagcctctactctactctatacagaatgtcgt 1560
 DB 1501 cgtctcgtcgtcgt 1560
 OY 1561 atgtcgtgaanaaaccaagacccataaagaatatagtctatgtcctgtgtgtgtgtgtgtgt 1620
 DB 1561 gttaacgtcgt 1620
 OY 1621 atggtgtgtatgtgaagcgtatgtacta 1646
 DB 1621 atggtgtgtatgtgaagcgtatgtacta 1646

RESULT 13

AAA48498
 ID AAA48498 standard; DNA; 1647 BP.

AAA48498;

07-DEC-2000 (first entry)

E. coli groEL coding sequence.

E. coli: groEL; proOmpA; trigger factor; outer membrane protein OmpA;

expression plasmid; ds.

Escherichia coli.

| XX | Key | Location/Qualifiers |
|----|-----|----------------------------|
| PH | CDS | 1..1647 |
| FT | | /*tag= a |
| FT | | /product= "GroEL" |
| XX | | |
| PN | | EP1016724-A2. |
| XX | | |
| PD | | 05-JUL-2000. |
| XX | | |
| PE | | 28-DEC-1999; 99EP-0126094. |
| XX | | |
| PR | | 28-DEC-1998; 98JP-0372965. |
| XX | | |
| PA | | (HSPR-) HSP RES INST INC. |
| XX | | |
| PI | | Sogo K, Yanaql H, Yura T; |
| XX | | |
| XX | | WPI; 2000-414804/36. |
| DR | | p-PSDB; AAY99621. |
| XX | | |

Novel operon comprising genes encoding a trigger factor, GroEL and GroES, useful for recombinantly producing foreign proteins.

Disclosure; Page 15-16; 24pp; English.

The present sequence is the coding sequence of groEL of *Escherichia coli*.

groEL is a protein which is involved in the folding of proteins, along with GROES. The genes for these two proteins can be used, along with the gene for a trigger factor which transports proompa to the outer membrane, to create an expression vector which is able to produce foreign

proteins in an insoluble form, which do not aggregate within the bacterial cell, but which have a stable conformation. The proteins capable of being

receptors, interleukin receptor agonists, granulocyte and macrophage and

granulocyte colony-stimulating factors, erythropoietin, thrombopoietin, leukaemia inhibitory factors, stem cell growth factors, tumour necrosis factor, growth hormone, and insulin-like growth factor-1.

[illegible]

inhibitors: neurokinase, tissue plasminogen activator, blood

coagulation factors, protein C, glucocerebrosidase, superoxide dismutase, renin, lysozyme, p450 prothymosin, trypsin and lactate dehydrogenase.

components, serum albumin, cedar pollen allergens, lipocortin, reptin, immunoglobulins, single-chain antibodies, complement, prostaglandin synthase, phospholipase, trypsin and elastase inhibitors, histamine, and leukotrienes.

hypoxia-induced stress
cyclic poliovirus antigens,
cyclic poliovirus antigens,
proteins, protein kinases,
proto-oncogene products,
regulation factors and virus-constituent proteins
transcription

Sequence 1647 BP: 437 A: 394 C: 476 G: 340 T: 0 other:
----- and ---- complementary proteins.

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

| | | | | |
|---------------------|--------|---------------------|--------|--------------|
| Query Match | 41.58; | Score 682.8; | DB 21; | Length 1647; |
| at Local Similarity | 63.48; | Pred. No. 4 5e-147; | | |

Matches 1044; Conservative 0; Mismatches 602; Indels 0; Gaps 0

1 atggtctctaagaatccttlltgatgctaaagcccgtagaaaacttcaagagtgta 60

1 atgcaagcgaagacgtataatctgtaacgaagctcgtgtgtaaaatgtctgcgcgcgtta 60

61 gataaacttgcaaatgctgttaagtaacacttgcacctaagcgcgtaatgtcgttatt 120

61 aacgtactggcagatgcagtgaaagttaacctcggtccaaaaagccgtaacgtagttctg 120

121 gaaagccttttgggtccccaagttaatacaaaaagaaggtytatcctglttgcaaaaaaat **180**

121 gataaattcttctcggtgcacccyaccatcaaccaagaatgtgtttccgttctgtctgtgaatc 180

101 gaacttgaagatacaatttgaataatattggtcgtccaatgtgttaagaagtagctcccaa 240

161 gaacgcggaagcacaagcttcgaaaaatattggtgcgcacgatgctgaagaagcttcctcctaaa 240

[illegible]

Db 1321 aaagttgcactgctgcaatggaagcctccgtgctcagatcgtatgaactgcgcgaa 1380
Qy 1381 gaagttctattgttgaagaaagatcgtgaacaaagatggtttgattatgct 1440
Db 1381 gaaacgtctgtctgttgaacacccgttaaaagcgcgacgaactaagcttaacacgca 1440
Qy 1441 gcaacgaagatatagaagacatttaaaagctgtgcatgatacctcaaaaattaca 1500
Db 1441 gcaacgaagatatacgcaacatgatacgtgatacctcgtgatacacaacaaatgact 1500
Qy 1501 cgtattgcatcaaaaatgacgacatagctctccttacttacttaacagaaatgcgc~1560
Db 1501 cgtctgctcgcaglaacgagctctgtgctgctgcatgatacacaacgaaatgcatg 1560
Qy 1561 atctcgtgaaaaccagaaacccaataaagatatgctatgctcgtgcggtggtatggtgt 1620
Db 1561 gtaccgacgtcgcgaaacgaatgcagctgacttaagcgtcgtcgtgatatggtgcgc 1620
1621 atgggtggtatggacgtatgacta 1646
1621 atgggtggtatggacgtatgacta 1646
Db 1621 atgggtggtatggacgtatgacta 1646

RESULT 14

AAA48500

ID AAA48500 standard; DNA: 4524 BP.

XX AAA48500;

XX 07-DEC-2000 (first entry)

XX Escherichia coli groES, groEL and tig operon coding sequence.

KW E. coli; groES; groEL; tig; proOmpA; trigger factor;
KW outer membrane protein OmpA; expression plasmid; ds.

XX Escherichia coli.

XX EPI016724-A2.

XX 05-JUL-2000.

XX 28-DEC-1999; 99EP-0126094.

XX 28-DEC-1998; 98JP-0372965.

XX (HSPR-) HSP RES INST INC.

XX Sogo K, Yanagi H, Yura T;

XX WPI: 2000-414804/36.

XX Novel operon comprising genes encoding a trigger factor, GroEL and
XX groES, useful for recombinantly producing foreign proteins -

XX Disclosure: Page 16-19; 24pp; English.

XX The present sequence is the coding sequence of an artificial operon
XX comprising the groES, groEL and trigger factor genes of Escherichia coli.
XX groES is a protein which is involved in the folding of proteins, along
XX with groEL. The trigger factor is involved in the transportation of
XX proOmpA to the outer membrane. The operon can be used to create an
XX expression vector which is able to produce foreign proteins in a soluble
XX form, which do not aggregate within the bacterial cell, but which have a
XX stable conformation. The proteins capable of being produced in this way
XX include interferons, interleukins, interleukin receptors, interleukin
XX receptor agonists, granulocyte, macrophage and granulocyte macrophage
XX colony-stimulating factors, erythropoietin, thrombopoietin, leukaemia
XX inhibitory factors, stem cell growth factors, tumour necrosis factors,
XX growth hormones, prolactin, insulin-like, fibroblast, platelet-derived,
XX transforming, hepatocyte and nerve growth factors, osteogenic, ciliary
XX neurotrophic, brain-derived neurotrophic and glial-derived neurotrophic
XX factors, neurotrophine, angiogenesis inhibitors, pronoxkinase, tissue

CC plasminogen activators, blood coagulation factors, protein C,
CC glucocorticosteroids, superoxide dismutase, renin, lysozyme, p450,
CC prothymosin, trypsin and elastase inhibitors, lipocortin, reptin,
CC immunoglobulins, single-chain antibodies, complement components, serum
CC albumin, cedar pollen allergens, hypoxia-induced stress proteins, protein
CC kinases, proto-oncogene products, transcription regulation factors and
CC virus-constituent proteins.

XX Sequence 4524 BP; 1229 A; 1058 C; 1246 G; 991 T; 0 other:

Query Match 41.5%; Score 682.8; DB 21; Length 4524;
Best Local Similarity 63.4%; Pred. No. 6.1e-147;
Matches 1044; Conservative 0; Mismatches 602; Indels 0; Gaps 0;

Qy 1 atgctcttaagaatccttctgtatgctaaagcccgatgaacatttcacgagtgta 60
Db 383 atgacgactaaagaacglaaaatctcgtaacgacgctgtgtgaatgctgcgcgta 442
Qy 61 gataaacttgcaaatgctgttaagtaaacacttggaacctaagggcgtatgtctat 120
Db 443 aacgtacgtgcagatgcagtgtaaaagttaaccctcgtgtccaaaagccgtaaacgtatctgc 502
Qy 121 gaaaagctcttgcgtcccaagttatacaaaaagatggtatctcgtgtgcaaaaagaatt 180
Db 503 gataaactcttcgtgcacacgacacatcaacaaagatggtgttcgtgtcgtgtaaatc 562
Qy 181 gaacttgaaagataagttgaaataatggcgctcaaatggttaaaagaaagtagctcccaa 240
Db 563 gaactggaagaaactgcgaataatagtggtgcagatggtgtaaaagattgctccctaaa 622
Qy 241 actagcgaatattgctggtggtatggaactacaagaacaaagctcttcgcaaaagcttatt 300
Db 623 gcaaaacgcgcgcgagcgagcgatgcgtacacccactgcgaacgcgtatgctatgctatcac 682
Qy 301 cgtgaaggtgtaaaactcgtgaacagctgtgcttaactatgcatcaaaagctgacata 360
Db 683 actgaaggtctgaagacgctgtgtcgtgcgtacgaagacccgagtgacctgaaagctgtgatac 742
Qy 361 gataaagctgtgtgtgtgttaactaaagaactaagcagacattacaagcctactcgtgac 420
Db 743 gacaaagcgtgtacccgctgcaagttgaagaacgtgaagacgctgttcgtaacagctctgac 802
Qy 421 caaaaagaatagctcaaggttggaacatcttcgtgaacactgatacaacaaatagtaatt 480
Db 803 tctaaagatgtgcaggttgcgtacacatcttcgtgaacactgcgaacacgtatgataaa 862
Qy 481 atcaatagctgaagctatgctaaagtgtgaaagaaggtgtatcacagtttgaaggaagct 540
Db 863 ctgatacgtcgaagcagatgtaacaaagctggttaagaagcgtatatacctgtgaagacggt 922
Qy 541 aaaggtctgaaaactaactatgatagtgtgtgtaagaagaagatttgaccgtgtgtaacctc 600
Db 922 accggtctgcgaaggaacatgagcgtgtgtgaaggtatgcaagctgtgtaacctg 982
Qy 601 tctcaatctgttaactaatctcgtgaagaatggttgcgtgaacttgataacactatc 660
Db 983 tctccttactatcaacaacgacggaacatggtgcgcatgaagacttgaagccgttcatc 1042
Qy 661 ctgttgaagagaaaagattactagcatgaagaacatgctacacatcttgaagaagtt 720
Db 1043 ctgcgtgcgtgaagaagaataatctcccaactcgcggaatgctgcggttcttggaagctgt 1102
Qy 721 gctaaagtaaacgcgtccactcttatattatgctggaagacgtgaagaaggtgaagcaacttga 780
Db 1103 gcaaaagcgcgaacacgcgtcgtcatcagctgtaagatgtaagaagcgaacgcgtgcga 1162
Qy 781 aacattgtagtaataagctcgtggaagaccccaagttgtagccgtgaataacatctggt 840
Db 1163 actcgtgtgtttaaaccatcttcgtgcaatcgtggaagctgcgtggtttaaaccacccgggc 1222
Qy 841 ctgtgtaacgcgttaaaactatgcttgaagatatgtcatcttacttgaagaagaaga 900


```

Db 1575 gttaacggacccgcgcgaagcgcatgcacccctgacttaagtgcctgcctgggtgatatggcgc 1634
QY 1621 atgggtgtgtatcg 1632
    ||||| || |||
Db 1635 atgggcgggaatg 1646

```

Search completed: August 1, 2001, 16:48:42
Job time: 4148 sec

Thu Aug 2 08:04:50 2001

us-09-077-574a-1.rng

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2001, 15:29:39 ; Search time 86.31 Seconds
(without alignments)
3612.506 Million cell updates/sec

Title: US-09-077-574a-1

Perfect score: 1647
Sequence: 1 atggtcttaagaagaalctt.....gtatggacgtatgtagtag 1647

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

arched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing firsts 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2_6/plodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/plodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/plodata/2/1na/6A.COMB.seq:*
4: /cgn2_6/plodata/2/1na/6B.COMB.seq:*
5: /cgn2_6/plodata/2/1na/PCTUS.COMB.seq:*
6: /cgn2_6/plodata/2/1na/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--|
| 1 | 682.8 | 41.5 | 1647 | 4 | US-09-472-971-5
Sequence 5, Appli |
| 2 | 682.8 | 41.5 | 4524 | 4 | US-09-472-971-7
Sequence 7, Appli |
| 3 | 642.4 | 39.0 | 1838 | 3 | US-08-470-260-7
Sequence 7, Appli |
| 4 | 642.4 | 39.0 | 1838 | 3 | US-08-471-491-7
Sequence 7, Appli |
| 5 | 642.4 | 39.0 | 1838 | 3 | US-08-466-662-7
Sequence 7, Appli |
| 6 | 610.4 | 37.1 | 2284 | 2 | US-08-467-822-28
Sequence 28, Appli |
| 7 | 610.4 | 37.1 | 2284 | 4 | US-08-432-697-28
Sequence 28, Appli |
| 8 | 396.6 | 24.1 | 1620 | 2 | US-08-461-775-10
Sequence 10, Appli |
| 9 | 396.6 | 24.1 | 1620 | 2 | US-09-031-606-10
Sequence 10, Appli |
| 10 | 390.8 | 23.7 | 1569 | 2 | US-08-997-080-113
Sequence 113, App |
| 11 | 390.8 | 23.7 | 1569 | 2 | US-08-997-362-113
Sequence 113, App |
| 12 | 390.8 | 23.7 | 1569 | 4 | US-09-095-855-113
Sequence 113, App |
| 13 | 390.8 | 23.7 | 1626 | 2 | US-08-997-080-159
Sequence 159, App |
| 14 | 390.8 | 23.7 | 1626 | 2 | US-08-997-362-159
Sequence 159, App |
| 15 | 390.8 | 23.7 | 1626 | 4 | US-09-095-855-159
Sequence 159, App |
| 16 | 390.2 | 23.7 | 2668 | 2 | US-08-461-775-11
Sequence 11, Appli |
| 17 | 390.2 | 23.7 | 2668 | 3 | US-09-031-606-11
Sequence 11, Appli |
| 18 | 344 | 20.9 | 1320 | 2 | US-08-461-775-8
Sequence 8, Appli |
| 19 | 344 | 20.9 | 1320 | 3 | US-09-031-606-8
Sequence 8, Appli |
| 20 | 334.4 | 20.3 | 2167 | 2 | US-08-461-775-9
Sequence 9, Appli |
| 21 | 334.4 | 20.3 | 2167 | 3 | US-09-031-606-9
Sequence 9, Appli |
| 22 | 225.6 | 13.7 | 927 | 2 | US-08-997-080-116
Sequence 116, App |
| 23 | 225.6 | 13.7 | 927 | 2 | US-08-997-362-116
Sequence 116, App |
| 24 | 225.6 | 13.7 | 927 | 4 | US-09-095-855-116
Sequence 116, App |
| 25 | 225.6 | 13.7 | 985 | 2 | US-08-997-080-161
Sequence 161, App |
| 26 | 225.6 | 13.7 | 985 | 2 | US-08-997-362-161
Sequence 161, App |
| 27 | 225.6 | 13.7 | 985 | 4 | US-09-095-855-161
Sequence 161, App |

| | | | | | |
|----|-------|------|------|---|--|
| 28 | 167.6 | 10.2 | 647 | 2 | US-08-997-080-115
Sequence 115, App |
| 29 | 167.6 | 10.2 | 647 | 2 | US-08-997-362-115
Sequence 115, App |
| 30 | 167.6 | 10.2 | 647 | 4 | US-09-095-855-115
Sequence 115, App |
| 31 | 137.2 | 8.3 | 568 | 3 | US-08-714-918-51
Sequence 51, Appli |
| 32 | 137.2 | 8.3 | 568 | 4 | US-09-265-315-51
Sequence 51, Appli |
| 33 | 137.2 | 8.3 | 568 | 4 | US-09-265-315-51
Sequence 51, Appli |
| 34 | 137.2 | 8.3 | 568 | 4 | US-09-266-417-51
Sequence 51, Appli |
| 35 | 93.6 | 5.7 | 360 | 2 | US-08-997-080-79
Sequence 79, Appli |
| 36 | 93.6 | 5.7 | 360 | 2 | US-08-997-362-79
Sequence 79, Appli |
| 37 | 93.6 | 5.7 | 360 | 4 | US-08-873-970-79
Sequence 79, Appli |
| 38 | 93.6 | 5.7 | 360 | 4 | US-09-095-855-79
Sequence 79, Appli |
| 39 | 89.6 | 5.4 | 337 | 2 | US-08-997-080-77
Sequence 77, Appli |
| 40 | 89.6 | 5.4 | 337 | 3 | US-08-997-362-77
Sequence 77, Appli |
| 41 | 89.6 | 5.4 | 337 | 3 | US-08-873-970-77
Sequence 77, Appli |
| 42 | 89.6 | 5.4 | 337 | 4 | US-09-095-855-77
Sequence 77, Appli |
| 43 | 86 | 5.2 | 343 | 2 | US-08-473-020A-2
Sequence 2, Appli |
| 44 | 85.8 | 5.2 | 1028 | 1 | US-08-029-328-1
Sequence 1, Appli |
| 45 | 85.2 | 5.2 | 343 | 2 | US-08-473-020A-8
Sequence 8, Appli |

ALIGNMENTS

| | | |
|--|--------------------|-----------------------------------|
| RESULT 1 | US-09-472-971-5 | Application US/09472971 |
| Sequence 5, Appli | Patent No. 6197547 | |
| GENERAL INFORMATION: | | |
| APPLICANT: SOGO, Kazuyo | | |
| APPLICANT: YAMAGI, Hideki | | |
| APPLICANT: YURA, Takashi | | |
| TITLE OF INVENTION: TRIGGER FACTOR EXPRESSION PLASMIDS | | |
| FILE REFERENCE: 1422-409P | | |
| CURRENT APPLICATION NUMBER: US/09/472, 971 | | |
| EARLIER FILING DATE: 1999-12-28 | | |
| EARLIER APPLICATION NUMBER: JP10-372965 | | |
| NUMBER OF SEQ ID NOS: 7 | | |
| SOFTWARE: PatentIn Ver. 2.1 | | |
| SEQ ID NO: 5 | | |
| LENGTH: 1647 | | |
| TYPE: DNA | | |
| ORGANISM: Escherichia coli | | |
| US-09-472-971-5 | | |
| Query Match | 41.5% | Score 682.8; DB 4; Length 1647; |
| Best Local Similarity | 63.4% | Pred. No. 3.2e-171; |
| Matches 1044; Conservative | 0; | Mismatches 602; Indels 0; Gaps 0; |
| QY 1 atggtcttaagaagaatccttttgatgcttaagcccgtagaaaactttcacgagtgta 60 | | |
| DB 1 atggtcttaagaagaatccttttgatgcttaagcccgtagaaaactttcacgagtgta 60 | | |
| QY 61 gataactgcaagcagctgttaagtaaacacttggacacccaagagccgtaatcgttat 120 | | |
| DB 61 aacgtactgacagcagctgttaagtaaacacttggacacccaagagccgtaatcgttat 120 | | |
| QY 121 gaaagctcttgggtcccaagttatataaagaatgtgtatctgtgtcaaaagaat 180 | | |
| DB 121 gataactcttgggtcccaagttatataaagaatgtgtatctgtgtcaaaagaat 180 | | |
| QY 181 gaacttgaagaataatgtgaataatgtggcgtcaaatgtgtaagaagtatgccaa 240 | | |
| DB 181 gaacttgaagaataatgtgaataatgtggcgtcaaatgtgtaagaagtatgccaa 240 | | |
| QY 241 actagcatatgtcgtgtgatggaactacaagaacagctccttgacaagaactttat 300 | | |
| DB 241 gcaaacgagcgtcgtgtgatggaactacaagaacagctccttgacaagaactttat 300 | | |
| QY 301 cgtgaagtgtaaaactgttagcagctgtgtcgttaatcctatgtgcaataaagctgata 360 | | |
| DB 301 actgaagtgtaaaactgttagcagctgtgtcgttaatcctatgtgcaataaagctgata 360 | | |

| | | | |
|----|------|---|------|
| OY | 361 | gataaagcgtctgttgcgtctgttactaaagaactaagcgacattacaagaactactcgtgac | 420 |
| Dp | 361 | gacaaacgcggcttcccgctctgcagtttgaaagacatgaaagcgtctccgtatccatgactctac | 420 |
| OY | 421 | caaaagaagaaatactcaagtttggaacctcttcgcgaactctgatacaacaatagttat | 480 |
| Dp | 421 | tcataagcgatgctccggtctgtgtaccactctccgcctaaactccgaagaacccgtatgtaa | 480 |
| OY | 481 | atcaaaagcgtaaagctaaagctaaagcttgaaagaaagagaggttatccacagtttgaaagaaagc | 540 |
| Dp | 481 | ctgactcgtgaagcgtatgacgaacaagctcgttgaaagaagaggttatccacgttgaaagaagcgt | 540 |
| OY | 541 | aaaggtcttgaaactaatatagttggtgttgaaagaaatgaatttgaccgggtgctacttc | 600 |
| Dp | 541 | accggtctcgaagagcaaacatggaacgtggttttgaaagtatgtaacattccgaccgttgcatactg | 600 |
| OY | 601 | tcctcaactcttgtaactaaacctcgtgaaagaaatggttttgaaactctgatacaaccttatctc | 660 |
| Dp | 601 | tcctcttatctatcaacaagaccgggaactggtcgcaagtgtgaactgtgaagcccggtctatc | 660 |
| OY | 661 | cttgtaatgagaanaagatctacatgacatgaaagacatgataccaactctgaacaagt | 720 |
| Dp | 661 | ctgtgtgctgtacagaagaaatctccaacacacccgaaatgctctccggtctctgaaagctgtgt | 720 |
| OY | 721 | gctaaatgaacgcgtcccaactcttatcttggtgaaagaaatgataagaggtgaaagcactctg | 780 |
| Dp | 721 | gccaaagcaggaacaaacgcgtcgtatcatalcgtctgaaagaatgtaaaagcgaagcgtctgca | 780 |
| OY | 781 | acactgttgatcataataagctccgctggagacatcccaagtttgtaagccgtataaagctctgtg | 840 |
| Dp | 781 | actcgtctgttaaccactctgtgtgcacatcgtgaaagatcgtcgtgttgtaaagcaccggcg | 840 |
| OY | 841 | cttggtgaaacgcgttgaaagatgacttgaaagatatgtcatcctctaacgtggaagaaagca | 900 |
| Dp | 841 | ctcgtcgatctcgttataaagatcatatgctgacgtgatatctgcgaacccctgacgtgcgtacgtg | 900 |
| OY | 901 | atacttgaaagatcgtgttgaaagagcttgaaagtgtgaagcttctctcttgtaagaaact | 960 |
| Dp | 901 | atctcgtgaagagatcgttatgtagagctcgtgaaagaaagcaaccctgtgaaagacctgtgtacagct | 960 |
| OY | 961 | aaacgctgactatctgacaaagaanaatactactactcgttatgactgtgtcgtgaaatacgaag | 1020 |
| Dp | 961 | aaacgcttcttgatacaacaagaacacacactatcatctgaatgagcgtgtggtgaaagaaact | 1020 |
| OY | 1021 | gatactaaagctcgaagttgaacaacatctgtgcaccaatttgaaagaaagaagaagcccaagat | 1080 |
| Dp | 1021 | gcaatccagagcgctgtgtctcgcagatccgtctcagcagatctgaaagaagaacactctgcatac | 1080 |
| OY | 1081 | gactctgaaanaactctcaagaacgctctcttgaaaaactgtgtgtgtgtagtgatttccat | 1140 |
| Dp | 1081 | gaccgttgaaanaactctcagagaacgctgtagcgaacatctgcagcgcgtgtgacgttatcaaa | 1140 |
| OY | 1141 | gttggtgaactctctcgtgaacttgaaatgaaagaaagaaagaaagatcgtgtgtaagaatgtctta | 1200 |
| Dp | 1141 | gtcgtgtgtgtgtctccgaaagttctgaaatgaaagaaagaaagaaagcgcgtctgaaagatgacctg | 1200 |
| OY | 1201 | aaatgcaaacaaagaaactcgtggtctgaaagaagatgactgtccctcgtgtgtgtgtaactcttctc | 1260 |
| Dp | 1201 | caagcgtgacccgctgtgtcgtgtgaaagaagagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt | 1260 |
| OY | 1261 | cgctccatctaaagttcctcttgatgatatataaacctgcgtgatgatagaaactgtgtgaact | 1320 |
| Dp | 1261 | cgctccatcgcttcaaaactcgtgtcgtacgcgtgtgtcagaaagaagaccagaacgtgtgtgtatc | 1320 |
| OY | 1321 | aaatctatccgtctctctcttgaaagagccttgaaacttgcaaatgtcgtcaaatgtcgtcat | 1380 |
| Dp | 1321 | aaagttgtcactcgtgtgaaatgaaagatcgcgtcgtcagatctgatatgaacacgcgtgcga | 1380 |
| OY | 1381 | gaaggtctctatgttgttgaaanaagttcgttgaaaccaaaaagatggtcttgatttaaagct | 1440 |
| Dp | 1381 | gaacagctgttgtgtgtgtaacaacgttgaaaggtgagcgtgacacgtgatacgtttataacgca | 1440 |

```

QY 1441 gcatcagaggaatatgaagaccttatttaaagctgtgttcattgatccctcaaaaagtcca 1500
Db 1441 gcaaccggaagaatacgcgaacacatgatcgacatgggattctcgtatcccaaccaaagtaact 1500
QY 1501 cgtattgcatcacaataatgcagcatcagtagcctctactcttaactacaagaatgcgt 1560
Db 1501 cgtctcgtctgcagtaacgcagctctctgtgtgcctgtatgatcacaccgaatgcatg 1560
QY 1561 attgcttgaanaaccggaaccttaaaaaagataatgcctatgcctgtgcgggttgatgggtggt 1620
Db 1561 gttaccgacgcctgcgaaaaaacgatcagctgcagcttaaggcgcctgcctggcggtatggcggtc 1620
QY 1621 atgggttgatgcagcgatgacta 1646
Db 1621 atgggtgcagatggcgcgcatgacta 1646

RESULT 2.
US-09-472-971-7
: Sequence 7, Application US/09472971
: Patent No. 6197547
: GENERAL INFORMATION:
: APPLICANT: SOGO, Kazuyo
: APPLICANT: YANAGI, Hideki
: APPLICANT: YURA, Takashi
: TITLE OF INVENTION: TRIGGER FACTOR EXPRESSION PLASMIDS
: FILE REFERENCE: 1422-409P
: CURRENT APPLICATION NUMBER: 1999-12-28
: EARLIER FILING DATE: 1999-12-28
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 7
: LENGTH: 4524
: TYPE: DNA
: ORGANISM: Escherichia coli
US-09-472-971-7

Query Match 41.5%, Score 682.8; DB 4; Length 4524;
Best Local Similarity 63.4%; Pred. No. 5.3e-171;
Matches 1044; Conservative 0; Mismatches 602; Indels 0; Gaps 0;

QY 1 atggtctctaaagaatcctcttttgatgctaaagccggtgaanaaacttcaagagtga 60
Db 383 atggcagctaaagaacgataaactctcgtaacgacgctcgtgtgaaaatgctgcggcgta 442
QY 61 gataaacttgcacaacgctgttaagaatlaaacacttgcacctaagaagccgtaatctgtatt 120
Db 443 aacgacgtacgcagatgcagatgaagtgaagttaacccctcgltccaaaagccgtaaactgttcg 502
QY 121 gaaaagctcttttgcttcccaagtattatcaaaagaatgtgtatctcgtttgcaaaagaatt 180
Db 503 gataaactcttcgtgcacccgcacacccaagaagaatgtgttccgttccgttcgtcgtaaatc 562
QY 181 gaacttgaagataaagtgtgaanaatatgtggcgctcaaatatgattgaagaagaatagctccaaa 240
Db 563 gaacttgaagaagaagttcgaanaatatgtgtgcgcagaatgtgtgaagaagaatgtgcctcaaa 622
QY 241 actagcgtatatctcgtgtgatatgaacctcaacaacagacagctcccttgcacaacgtatttat 300
Db 623 gaaacagcagcgtcgaagcgacggtacacacacttgcacaacgctatcgtctcaggtcatatcac 682
QY 301 cgtgaaggtgtlaaaaactcgttagcagctgtgtcgttaactcctatgagcaattaaacgttgcata 360
Db 683 actgaaggtgtcgtgaagcgtgtgtcgtcgtgcgtacgtgaacccgattgacctgtgaacgttgatc 742
QY 361 gattaagcgtgtgtgtcgttacttaaaagactaaagcgaatcgtaatcaaaagctctcgttgc 420
Db 743 gacaagcgtgttaccgcgtgcagatgtgaagaacttgaagaagcgtgttccgtlaccatgctctgac 802
QY 421 caaaaagaatatgcctcaagtttgaaccatcttctgcaaacctcgtatatacaacaatagtaat 480

```

```

Db      803  tctaaagcgattgctcaggttggtacccatcccgtaactccgaagaaacgtaagtaa 862
Oy      481  atcatagctgaagctatgagctaaagtlgaaaagagaggtgtatcacagttgagaagct 540
Db      863  ctgaatcgctgaagcgtatgacaaagctcgtaaaagagcgctatcacgcttgaaagcgt 922
Oy      541  aaagctctgaactacatagatctgtctgaagaaagaaagtttgaccgtgctaccc 600
Db      923  accggtctgcagagacgaactcgtgacgtgttgaaagttatgcagcttcgaccgtgctac 982
Oy      601  tctccatacttctaactaatccctgagaaaatgtgttgaaactgtataacccttatac 660
Db      983  tctcctacttcatcaacaagcggaaaactggtgcagtagaactgtgaagccgcttacc 1042
Oy      661  ctctgtatagaaaagatatctactagcttgaaagacatgctaccacttgaacaagt 720
Db      1043  ctgctgtgcgaaagaaaatctccacacatccggaatgctgcccgttcttgaaagctgt 1102
Oy      721  gctaaagtaaacgcctccatctcttattgtctgaagacgtagaagctgaaacattgca 780
Db      1103  gccaaagcagagcaaacgcctctgatactcgtctgaagatgtatgaagcgaaagcgtgtg 1162
Oy      781  acaactgtgatacaataagctccgtggaagcactccaagtgtgaagccgtataaagctcgt 840
Db      1163  actgctgtgttaacacacatctggtgcaatcgtaaaagtcgctgcgtttaaagcaccggt 1222
Oy      841  ttgtgtgaacgcgttaaaagctatgcttgaagatatctacttaccgtgaagagaagca 900
Db      1223  ttcggtgatacgtctaaagcgtatctgtagagatactgaacacccctgacgtgcgtacgt 1282
Oy      901  atatttgaagatcgtgtgataaagcttgaagcttgaagcttctctcttggaacagct 960
Db      1283  atctctgaagagatcgtatgagatcgtgaaaagcaaaccccggaagacctggtgtcagct 1342
Oy      961  aaacgtgtgattatctgaacaaagaaatactactcgttgaatgtgtcgtgaagaaatcaga 1020
Db      1343  aaacgtgtgtgatacaacaagaaacacacacatctatcgtatgtgtgtgtgaaagagct 1402
Oy      1021  gatatataagctcgaatlaaaccaaatctgtgcacaaatctgaagaagaacagctcagat 1080
Db      1403  gcaatccagggcgtgtgtgcacagatccgtcagcagatctgaagaagcaactcttgacac 1462
Oy      1081  gatcgtgaagaaacttcaagaaacgtctctgaacaaactgtgtgtgaagtgagctgtatc 1140
Db      1463  gaccgtgaaaaactgcaggaagcgtctgacgaagaaactgcgagcggtgttcgcttatac 1522
Oy      1141  gctgagctgtctactgaactgaactgaagaaagaaagagatcgtgtgaagaagatgctcta 1200
Db      1523  gtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1582
Oy      1201  aatgcacaagaagctgcgtgtgaaagagatgtgtccctgtgtgtgtgtgtgtgtgtgt 1260
Db      1583  caccgacccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1642
Oy      1261  cgtcccatlaaagctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1320
Db      1643  cgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1702
Oy      1321  aatatcatcgtctgtctctctgtgaagagccttactacatgctgaacatgctgtgctat 1380
Db      1703  aaagtgtgacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1762
Oy      1381  gaaggtctatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1440
Db      1763  gaacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1822
Oy      1441  gcatcaggaatactgaagacatctataaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1500
Db      1823  gcaacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1882
Oy      1501  cgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1560

```

```

Db      1883  cgtctgtctgcagtaacgaactctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1942
Oy      1561  attgtctgaaaaaacgaactctaaagataagctatgctcgtgtgtgtgtgtgtgtgtgt 1620
Db      1943  gtaccgacccgtgcgaaaaacgaactgaactgaactgaactgaactgaactgaactgaact 2002
Oy      1621  atggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1646
Db      2003  atggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2028

RESULT 3
US-08-470-260-7
: Sequence 7, Application US/08470260
: Patent No. 6077706
:
: GENERAL INFORMATION:
: APPLICANT: Covacci, Antonello
: APPLICANT: Bugnoli, Massimo
: APPLICANT: Telford, John
: APPLICANT: Macchia, Giovanni
: APPLICANT: Rappelli, Rino
:
: TITLE OF INVENTION: Helicobacter Pylori Proteins Useful
:
: NUMBER OF SEQUENCES: 7
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Chiron Corporation
: STREET: 4560 Horton Street
: CITY: Emeryville
: STATE: California
: COUNTRY: USA
: ZIP: 94608-2916
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/470,260
: FILING DATE:
:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/256,848
: FILING DATE: 21-OCT-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: McClung, Barbara G.
: REGISTRATION NUMBER: 33,113
: REFERENCE/DOCKET NUMBER: 0316,001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (510) 601-2708
: TELEFAX: (510) 655-3542
:
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1838 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: MOLECULE TYPE: DNA (genomic)
:
: US-08-470-260-7

Query Match      39.0%: Score 642.4; DB 3; Length 1838;
Best Local Similarity 62.9%: Pred. No. 1.6e-160;
Matches 1032: Conservative 0; Mismatches 601; Indels 9; Gaps 2;

```

```
Db 182 GCTATGGCGCTCAAGCATCACCAAGACGGCGTGAGGTGGCTAAAGATTAATTA 241
Qy 188 aagaatgatttgaataatgagctcaaatggttaagaagtgtccccaactgagc 247
Db 242 GTTCCCAAGTAGCTAACATGGGCGCTCAACTCGTTAAAGAGTGGCAGCAAAACCCCTG 301
Qy 248 atattgctggtatggaactaacacagcaacagctcttgcacaaagctattatgtag 307
Db 302 ATGCTGCCCGCGATGGACACGACCGACCGTACCTTATGATGATTTTAAAGAG 361
Qy 308 gtagaaacttgaacagctggtgctlaactctatgacatlaaacgctgagcaataag 367
Db 362 GTTGAAGAAATATCAGCGCTGGGCTAACCTTATGAAGTGAACGAGGATGATTAAG 421
Qy 368 ctgttctgtcttgaacaaagcaacatcacaagcctcctgtagccaaag 427
Db 422 CTGCTGAAGCGATCATTAATGAGCTTAAAGAGCAGCAAAAAGTGGCGGTAAAGAG 481
Qy 428 aatagctcaagttggaacacattctgcaactctgaacacatagtaatactag 487
Db 482 AAATCACCACCAAGTGGCCACCATTTCTCAAACTCCGATCACAAATATCGGAAACTCATCG 541
Qy 488 ctgaagctatggtctaaagttgaaagaaggtgctatcacagttgagaagctaaagtc 547
Db 542 CTGACGCTATGGAAGAAAGTGGTAAACAGCGCGTGAATCACCCTTGAGAGAGCTAAGGCA 601
Qy 548 ctgaactacaattagatggttgaagaatgaaagttgacccgtgctacccctccat 607
Db 602 TTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661
Qy 608 acttgaactaaaccgaaagaaatggttgaacttgaacttgaacttatacttctgta 667
Db 662 ATTTTGAACCAAGCGCTGAGAAATGACCCCTCATTTGATTAATGCTTTAATCTTTAA 721
Qy 668 atgagaaagatctactagatgaagaacalgtctacacatcttgaacaaagttgctaaag 727
Db 722 CGGTTAAATAAATCTTACATGAAGACATTTCCCGCTACTAGTAAAGAAACCACTGAAG 781
Qy 728 taaacgtccaccccttactatctgtagaagcgttagaaggtgaagcaactgcaacactg 787
Db 782 AGGCAACACCGCTTTTATATCTCCTGACATGAGGCGGAAAGCTTTAAACGACTAG 841
Qy 788 taatcaataagctccgtgagacacacacacacacacacacacacacacacacacacacac 847
Db 842 TGGTGAATTAATTAAGAGCGCTGTGAATATCCGACGGTTAAAGCTCCAGCGCTTGGG 901
Qy 848 aacgcgtlaaagctatgctgaagatactcactctactgagagagaaacatactg 907
Db 902 ACAGAAAGAAAGAAATGCTCAAGACATCCGCTATTAAACCGCGGCTCAAGTCTTTAGCG 961
Qy 908 aagatcgtgtataaagcttgaagaatgaaagctgtctctctttagaagcaagcgtg 967
Db 962 AAGAAATGGGCTTGTAGTGAAGAAACCGTCAAGTGGAGTTTAAAGCAAGCTGGAAGGA 1021
Qy 968 tagtattgacaaagaaatactactctgtagaagctgtagaagaaatcgaagatact 1027
Db 1022 TTTCGATTTGACAAAGACACACACGATCGTAGTGGCAAAAGCGCATAGCGATGATTA 1081
Qy 1028 aagctcagtttaaaccaattctgtagacaatactgaagaagaagctcagatlatatgctg 1087
Db 1082 AAGACAGATGCGCCACATCAAAACCAATTTGCAAGTACGACAGGATTTTGGACAAAG 1141
Qy 1088 aaaaactcaagaagcctctgcaaaactgttgaagtagtctgtatccatctgtagag 1147
Db 1142 AAAAATTTGCAAGAAAGATTTGGCTAAACTCTGCGCGGTGTGCTGATTAAGGTGGCG 1201
Qy 1148 ctgctactgaaactgaagaagaagaagaagatgctgtagaagatgctcclaagtcaa 1207
Db 1202 CTGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1261
Qy 1208 caaagagctgcgttgaagaagatgctcctgtagttagtcttctgtagcgcctca 1267
```

```
Db 1262 CTAAAGCGCGGTGAAGAGCATGTGATGTGTGCGGGTGGCGCTCATTCGCGCG 1321
Qy 1268 ttaagctccttgatgatacttaaacctctgtatgtagtaactgtgagcaatactca 1327
Db 1322 CTCAAAAAGTCAATTTGAATTTGC-----ACGATGATGAAGAAAGTGGCGCTATGAATTA 1375
Qy 1328 tccgtctgctcttgaagaagccttaagcttgaacttgaacttgaacttgaacttgaacttgaact 1387
Db 1376 TCATGCCCGCATTAAGAGCCCATTTAGCTCAAAAGCTGATCAACGCTGTTATGATGATGCGC 1435
Qy 1388 ctatgtctgaagaagaagctctgtagaagaagatggttctgtagaacttgaacttgaacttgaact 1447
Db 1436 GTGTGCTGCTGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1495
Qy 1448 gagaataatgaagccttataaagctggtgctatgtagcctcaaaagtagtaactgtag 1507
Db 1496 GCAAGTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1555
Qy 1508 catcacaaatgtagacacacacacacacacacacacacacacacacacacacacacacacacacacac 1567
Db 1556 CTCTACAAATATGCGGCTTCCGTTTCAAGCCCTGTTTAAACACAGAGCCACCGCTGATG 1615
Qy 1568 aaaaaccagaacacaaagataatgcttagctgctg---gagtgtagtggtagtagg 1624
Db 1616 AAATCAAGAAAGAAAGAAAGCACTCCGCGCATTCCTGATATGCTGATGCTGATGCTGATGCTGATGCTGAT 1675
Qy 1625 gtgtatgtagcgtatgtagta 1646
Db 1676 GAGGATGAGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1697

RESULT 4
US-08-471-491-7
: Sequence 7, Application US/08471491B
: Patent No. 6090611
: GENERAL INFORMATION:
: APPLICANT: Covacel, Antonello
: APPLICANT: Bugnoli, Massimo
: APPLICANT: Telford, John
: APPLICANT: Macchia, Giovanni
: APPLICANT: Rappelli, Rino
: TITLE OF INVENTION: Helicobacter Pylori Proteins useful For Vaccines And
: FILE REFERENCE: CHIN004
: CURRENT APPLICATION NUMBER: US/08/471,491B
: CURRENT FILING DATE: 1995-06-06
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: Patentin Ver. 2.1
: SEQ ID NO 7
: LENGTH: 1838
: TYPE: DNA
: ORGANISM: Helicobacter pylori
US-08-471-491-7

Query Match 39.0%; Score 642.4; DB 3; Length 1838;
Best Local Similarity 62.9%; Pred. No. 1,6e-160;
Matches 1032; Conservative 0; Mismatches 601; Indels 9; Gaps 2;
```

```

Db 242 gtgtccctgtagctacatctggcgtctcaactcgtctaaagaagtgcgagcaaacgcgtg 301
Qy 248 atattgcttgcgtgtagaactcaacaacagacagctctctgcaacaactctatctatcgtgaag 307
Db 302 atgtcgcgcggcgtatgtagcagcagcagcagcgtctgtagctatatacgtcttctttaaagag 361
Qy 308 gctgtaaaactctgtagcagcgtctgcttaactctctatgagccataaactcgtgcagatagaag 367
Db 362 gtttgaggatatacagcgtctggggcttaaccctatggaagtgaagcagagcattgtagaag 421
Qy 368 ctgtctgtctgctctactaaagaactaagcagacatctacaaagcctctctgtagccaaaag 427
Db 422 ctgtcgtgaagcgtatcatatgacttaaaagcgtgacaanaagctgaagcgtgtaagaag 481
Qy 428 aaatgctcaagcttggaaccattctctgcaactctgatacaacaataatgataatcatag 487
Db 482 aaatcacccaagctgagccattctctgcaaacctccgacatctacacatctcgggaaccatcag 541
Qy 488 ctgaagctatgctgctaaagtctgaaaaagaggtgtatcacagcttgaggaagcctaaagctc 547
Db 542 ctgacgctatggaanaaagtgggtaagaagcgtgtagacgcttgaggaagcctaaagc 601
Qy 548 ttgaacacatcatatgtagtctgtaagaagatggaagttgacccgtgtactctctcat 607
Db 602 ttgaagatgaattgtagtctgtagaagcgtgcaattctgataagagctacacctctccctc 661
Qy 608 acttgttaacttaactctctgtagaanaagtgtctgtaactctgtaacccttatactctctgta 667
Db 662 atttctgtaacgaacgctgtagaanaagtgtacccgtcctaattgataatctctacacctttaa 721
Qy 668 atgagaanaagatctctacatctgtagaagacatgctacccaactctgaagaagtctgtaag 727
Db 722 cgtgataaanaaactctctacatctgtagaagacatctctccgctactagaanaaacaactgaag 781
Qy 728 taacacgctccactctctatctgtctgtagaagcgtgtagaaggtggaagcactctgtaaac 787
Db 782 agggcaaacccgtcttctatctcgtctgtagaagacatctgtagaagcgtctttaaagccttag 841
Qy 788 tagtcaatlaagctctccgtgtagaagcctcccaagtgtgtagccgttaaaagcctctgttctgtg 847
Db 842 tggtagaataaataaagaagcgtgtgtgtagatctgcagcagcgtgtaaaagcctccagcgttgg 901
Qy 848 aacgcgtgaagctatgctgtagaagatctgctactctactctgtagaagagaaacatattctg 907
Db 902 acagaagaagaagaatgctcacaagacatctgctattcttaacggtcggtcgaagcattagcg 961
Qy 908 aagatcgtggtatgaagcttgaagaatgtgaagctgtctctctttagaagaagcctaaagctg 967
Db 962 aagaaatggtgctgtagcttagaagaacgctggaagtgaagttttagaagaagcctggaagga 1021
Qy 968 tagttatgacaagaagaataactactactctgtatgtagtgcgtggaagaatcagaagaatatta 1027
Db 1022 ttgtgtatgacaagaagaacacacacacagatcgttagatgtagcaaaagcctagcgtgtatgta 1081
Qy 1028 aagcgcgtgcttaaaacaaatctgtgcacaacatgtagaagaacaagcctcagaattatgctgtg 1087
Db 1082 aagacagaggtcgcgcagatcaaaacccaatctgacaagtacgacaagcagatattgcaaaag 1141
Qy 1088 aaaaactctcaagaacgctctgcaaaactgtgtgtgtagaagcgtttatccatgtgtgag 1147
Db 1142 aaaaactctcaagaagaagatgtgctaaactctcgtgcgtgtgtgtgtctgtatgaaggtggcg 1201
Qy 1148 ctgtctactgaaaactgaaaatgaaaagagaagaagatcgtgtgaagaatgctcttaaaatgca 1207
Db 1202 ctgtcagatgaaatgaaaatgaaaagagaagaagaacggtgtgtagatcagcgtttagacgcgca 1261
Qy 1208 caagcgtctgctgtgtagaagaagatctgtccctgctgtgtgtgtctctcttctgtccgtcca 1267
Db 1262 ctaaaagcgcgtgtgtagaagaagcattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1321
Qy 1268 ttaaaagctctgtagatataaactgtctgtagatgtagaagacatgtgtgtagatataatca 1327
Db 1322 ctcaaaaagctgtagatataaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1375

```

```

Qy 1328 tccgtctctctctgtagaagaagccttaacgtcaaatgtctgcaaatgtctgtagaaggtc 1387
Db 1376 tcatgtgcgcatttaagaagcccatctagctcaaatctgatacaagcgtgtatgtagtgcg 1435
Qy 1388 ctatgtctgtgtagaagaaggtctgtgaacccaagaagatgtgtgtgtgtgtgtgtgtgtgtgtgt 1447
Db 1436 gttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1495
Qy 1448 gagaatgtagaagccttaaaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1507
Db 1486 gcaatgtagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1555
Qy 1508 catcaaaaatgtagaagccttagcgtctctctctctctctctctctctctctctctctctctctct 1567
Db 1556 ctctcaaaaatgtagaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1615
Qy 1568 aaaaacgtagaagccttaaaaatgtagaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1624
Db 1616 aaatcaagaagaagaagaagcgtctcgtgcaatgtcgtgatagtgtgtgtgtgtgtgtgtgtgt 1675
Qy 1625 gtagtagtagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1666
Db 1676 gaggcagtagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1697

RESULT 5
US-08-466-662-7
; Sequence 7, Application US/08466662B
; Patent No. 6130059
; GENERAL INFORMATION:
; APPLICANT: Covaccl, Antonello
; APPLICANT: Bugnoli, Massimo
; APPLICANT: Telford, John
; APPLICANT: Macchia, Giovanni
; APPLICANT: Rappuoli, Rino
; TITLE OF INVENTION: Helicobacter pylori Proteins Useful For Vaccines And
; FILE REFERENCE: CHIR0057
; CURRENT APPLICATION NUMBER: US/08/466,662B
; CURRENT FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1838
; TYPE: DNA
; ORGANISM: Helicobacter pylori
US-08-466-662-7

Query Match 39.0%; Score 642.4; DB 3; Length 1838;
Best Local Similarity 62.9%; Pred. No. 1.6e-160;
Matches 1032; Conservative 0; Mismatches 601; Indels 9; Gaps 2;

```

| | | | |
|----|------|---|------|
| QY | 308 | g l g u a a a a c t t g t a g c a g c t g g l c g t a a t c c t a t g c c a t t a a c a g t g g c a t a g a a g | 367 |
| Db | 362 | g t t c g t a g a a t a c a c g c g t g g g c t t a a c c a c t a t t a a g t g a a g a c g a a c g a t g t a a a g | 421 |
| QY | 368 | c t g t g t g t g t t a c t a a a g a a c t a a g c a c t t a c a a g c t a c t c g t g c a a a a g | 427 |
| Db | 422 | c t g c t g a a g c g a t c a t t a t a g a g c t t a a a a a g c g a a a a a g t a g c g t a a g a a g | 481 |
| QY | 428 | a a a t a g c t a a g t t g a a c c a t t c t g c a a a c t t g t a t a c a a t a g t a t a t a t a g | 487 |
| Db | 482 | a a a t c a c c c a a g t g g c a c c a t t c t g c a a c t c g c a t c a a t t c t g g a a c t a c g | 541 |
| QY | 488 | c t g a a g c t a g c t a a a g t t g a a a a g g g g t t a c c a a t t g a a g a a g a a g t a a g t c | 547 |
| Db | 602 | t t g a a g t a g a t t g a t c g t a a a g a g c a t g t a a t t t g a t a g a g c t a c c t c c c t t | 661 |
| QY | 608 | a c t t g t a a c t a a t c c t g a a a a t g t t t g t a a c t t g t a a c c t a t a c t t a c t t t g t a | 667 |
| Db | 662 | a t t t t g a a c g a a c g t g t a g a a a a t g a c c g c t a a t t g t a t a a t g t t a c a c t t t t a a | 721 |
| QY | 668 | a t g a g a a a a a g a t t a c t a g c a t g a a a g a c a t g c t a c c a a t t c t t a a a c a a g t t c t a a g | 727 |
| Db | 722 | c g g t a a a a a a a t c t c t a g t a g t a a a g a a c a t t c c c g t a t a g a a a a a a c a t g a a a g | 781 |
| QY | 728 | t a a a c g t c a a c c t t a t a t t g t c g t a a g a c g t a a a a g t g a a c a c t t g c a a c a t g | 787 |
| Db | 782 | a g g g c a a a c c g c t t t a a t c a t c g c t g a a g a c a t t g a a g g c g a a g c t t t a a a g c t c a g | 841 |
| QY | 788 | t a g t c a a t a a g c t c g t g a g a c a c t c c a a g t c t g a g c g t a a a a g c c c c t g t t t g t g | 847 |
| Db | 842 | t g t g t a a t a a t t a a a g a g c g t g t t g a t a t a c g a c g g t t l a a a g c t c c a a g c t t t g g g | 901 |
| QY | 848 | a a c g c c g t a a g c t a g c t t g a a a g a t a t g c t a a c t t a c t t a g a g a a g a a t a t t g | 907 |
| Db | 902 | a c a g a a a a a a a a g a a c g a c c a a a g a c a t c g c a t t t a a c c g g c g a g t c a a g c a t t a c g | 961 |
| QY | 908 | a a a t c g t g t a t a a a g c t t g a a a t g t a a g c t g c t c t c t t a g a a c a g t l a a a c g t g | 967 |
| Db | 962 | a a g a a t t g g c t t g a g t c t a g a a a a c g c g a a a g t g a g t t c t t a g c a a a g c t g a a g a | 1021 |
| QY | 968 | t a g t t a t c t a a a a a a t a c t a c t a c c g t c g a t g t g t c g t g a a a a t c a a g a g a t a t a | 1027 |
| Db | 1022 | t t g t a t t g c a a a g a c a a c a c c a c a g a t c g t a g a t g c a a a g c a t a g c g a t g t a t a | 1081 |
| QY | 1028 | a a g c t g a a g t t a a c a a t t c g c g c a a a t t g a a a a a a a g c c c a g a t a t a g t t c g t | 1087 |
| Db | 1082 | a a g c a a g a g t c g c g c g a t c a a a a c c c a a a t t g c a a g t a c g a a a g c g a t t a t g c a a a g | 1141 |
| QY | 1088 | a a a a a c t t c a a a a c g t c t g c a a a a c t t g t t g t g a a t a g c t t a t c a t g t t g a a g | 1147 |
| Db | 1142 | a a a a a t t g c a a a a a g a t t g g c t a a a c t c t c t c g c g t g t g c g t g t a t a a a g g c g t | 1201 |
| QY | 1148 | c t g c t a c t g a a a c t g a a a t g a a a g a a g a a g a t c g t g t a g a a g a t g c t a a a t g c a a | 1207 |
| Db | 1202 | c t g c g a a t g a a g t g a a a t a t a a a g a a a a a a a a c c g g t t g a t a c a c g t t g a a c g a | 1261 |
| QY | 1208 | c a a a a g t g g g t t g a a a a g a g t a t g t c c c g g t g t g t a a c t g t t t t g c c g c t c a | 1267 |
| Db | 1262 | c t a a a g g g g g t t g a a a a g a g a c a t g t a t c t g t g g c g t g c g c t c c t a c t c g c g g | 1321 |
| QY | 1268 | t t a a a g t c t t g t a t a t a a a c c t g c t g a t a g t a g a a c t t g c t g a a t t a a t a c a | 1327 |
| Db | 1322 | c t c a a a a g t g a c t t g a a t t t g c ----- a c a g a t a t g a a a a g g g c t a t g a a c a | 1375 |
| QY | 1328 | t c a g t c t c t c t t g a a a g a c c t t t c g t c a a a t a t g c t c a a a t g c t g c t a t a a g t c | 1387 |
| Db | 1376 | t c a t a g c g c c a t t a a a g c c c a t t a g c t t a a a t c g c t a c a a g c t g t a t a t a g t a g c g | 1435 |

```

Oy 1308 ctttctgttgtaaaagaattcgttgaccacaaagaatgttlttgatattaaatcgtacag 1447
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1436 gtgtgtgtcgtgaatcgaagtagaaaaaacgaagggcatttctttaaagctagaatg 1495
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 1448 gagaataagaaaccattataaagctgtgtgtcattgtaacctaaaaagtacacgtatg 1507
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1496 gcaagtagtgcataatgtttaaagaagcattattgaccccttaaaagtagaaagatcg 1555
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 1508 cattacaaatgcagcatctagtcgcccttactctcaactaaagaatgagtcattgcg 1567
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1556 ctctacaaaatggcgttcttgcttcaagcctgtttaaaccacagaagcacccgtgcattg 1615
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 1568 aaaaaccgaaccacaaaagaataatgcctatgcctg--gcggtgatgtggtgatg 1624
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1616 aatcaacagaagaanaaagcgactccgcgcaatgcctgatgtgggtgcagtggcgcatg 1675
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 1625 gtgtgatgagacgtcatgtacta 1646
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1676 gaggcatgtgcgcgatgatgla 1697
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 6
: US-08-467-822-28
: Sequence 28, Application US/08467822
: Patent No. 5843460
: GENERAL INFORMATION:
: APPLICANT: Labigne, Agnes
: APPLICANT: Saueraum, Sebastien
: APPLICANT: Ferrero, Richard L.
: APPLICANT: Thiberge, Jean-Michel
: TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
: TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
: TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
: NUMBER OF SEQUENCES: 44
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
: ADDRESSEE: Dunner
: STREET: 1300 I Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/467,822
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/447,177
: FILING DATE: 19-MAY-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/432,697
: FILING DATE: 02-MAY-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Meyers, Kenneth J.
: REGISTRATION NUMBER: 25,146
: REFERENCE/DOCKET NUMBER: 03495.0137-02000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 408-4400
: TELEFAX: (202) 408-4400
: INFORMATION FOR SEQ ID NO: 28:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2284 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear

```


MOLECULE TYPE: DNA (genomic)
US-08-467-822-28

Query Match 37.18; Score 610.4; DB 2; Length 2284;
Best Local Similarity 62.18; Pred. No. 5.2e-152;
Matches 1019; Conservative 0; Mismatches 611; Indels 12; Gaps 3;

```
OY 8 ctgaagaalaccccttctgactgaagcccgctgaagaaccccttcacagagctgagtaaac 67
DB 510 CAAAGAAATCAAAATTTTCAGATAGCCGACAGAAACCTTTTATTGAGGCGTAAACAAAC 569
OY 68 ttgcaaatgctgtaagaatacaactgacctaagaagccgtaaatgctgtaactgaagaat 127
DB 570 TCATAGACGCTGTCAAACTATACATGGGCGCAAGGACGAGACGCTGTGATCCAAAAA 629
OY 128 ctttggtcccgatcttacaagaatgagtgatctgtaacgaagaatgaacttg 187
DB 630 GCTATGCGCGCTCAAGCATCACCAAGACGCGCTGAGCGCTGAAGAGATTTGAATTAA 689
OY 188 aagataaagcttgaagaataatgagccgctgaagaatgagtaagctcccaaaactagcg 247
DB 690 GTTGCCCGCTGCTACATGGGCGCTGAGCTGTTAAAGAGATGCGAAGAAACCGCTG 749
OY 248 aatgtgctgctgtaagaactacaagacagacagctctgcaagaactatctatcgtaag 307
DB 750 ATGCGCGCGGCGATGCGACGACGACGACGCGCTGCTGCTTATAGCATTTTAAAGAG 809
OY 308 gtagtaaaactgtagcagctgctgtaactcctatgacctaagaactgagcagataaga 367
DB 810 GCTTACGAAATATCACGCGCTGCGCTTACCTATTGAACGTGAACGACGATGTAAG 869
OY 368 ctgtgtgctgcttactaagaactaagaacgacatacaagaactcctgtaacccaaga 427
DB 870 CGCGTGAGCGATCTTATAGCTTAAAGAGCGAAGCAAAAGTGGCGGTTAAAGAG 929
OY 428 aatgtcgaagctgtaagaactcttgcgaactcgtatcaacaagaatgataatcatag 487
DB 930 AAATACCCAGTACGACGACATTTCTGCAACTCGCATACATATCGGAAACGCAATCG 989
OY 488 ctgaagctagctgaagaatggaagaagtgtagtaacagctggaagaagctgaagtc 547
DB 990 CTGACGCTATGGAAGAAATGCGCTTAAGACGCGCTGATCACCTGTAAGAGCTTAAG 1049
OY 548 ttgaacacatataatgctgtagtaagaatggaagctggaagctgtaacccctcctcat 607
DB 1050 TTGAAGATGAATTAAGATGCTGTAAGAGGATGCAATTTGATAGAGGCTACCTCCCTT 1109
OY 608 actgtgaactaactcctgagaagaatggtttgtaactgtaacccctatcactctgta 667
DB 1110 ACTTGTATACCAACGCTGAGAAATGACCGCTCAATTTGATATACGCTTACCTTTTAA 1169
OY 668 atgagaagaagaatctagatgataagaacatgctacccaactctgaagaagctgtaag 727
DB 1170 CGGATAAAAAATCTCTGATGAGAAAGACATTTCCCGCTACTAAAAAACAACATGAAG 1229
OY 728 taacagcgcactccttatatgctgtaagaagctgagaagtgtaagaactgtaacaactg 787
DB 1230 AGGGCAAAACCGCTTTAATCATGCTGAGAGACATTTGAGGCGCGAAGCTTTAAAGACTG 1289
OY 788 tagtaataaagctgtagagacaccccaagctgtagccgtaagaactcctgctgtagtg 847
DB 1290 TGGTGAATTAATTAAGAGCGCTGTTGAATATGCAAGCGTTAAAGCTTCAAGCTTTGGG 1349
OY 848 aagcgcgtgaagaactgctgtaagaatgctatcctccttaactggaaggaagaactatgtg 907
DB 1350 ACAGAGAGAAAGAAATGCTCAAGACATGCGTGTTTTAAACGCGGCTCAAGCATTAACG 1409
OY 908 aagatcgtagtgaagaactgtaagaatgtaagctgtctctcttaggaagaactgaagctg 967
DB 1410 AAGATTTGGGCTTGAAGTCAAGAAACGCTGAAGTGAAGTTTAAAGCAAAAGCA--GA 1466
OY 968 tagttatgacaagaataactactatcgtgtgtagtgctggaagaatcagaatatatta 1027
```

```
DB 1467 TTTGATTTGACAAAGACACACGATGATGATGGCAAGGCCATTAACCATACGCTCA 1536
OY 1028 aagcctgagcttaacaagaatctgtagcaaatggaagaagaagctcagatctatgctg 1087
DB 1527 AAGACAGAGTGGCGCAATTCAAAAACCCAAATTTGCAAGCAGCAAGCATTAACGAANG 1586
OY 1088 aaaaactcaagaagctcttgcgaagaactgctgtaggaagctgtagtctatccatgttgag 1147
DB 1587 AAAATTCAGAAAGATTTGGCGCAACCTCTCGGGCGGTGGCTGTGATTTAAAGTGGCG 1646
OY 1148 ctgctactgaagaactgtaagaagaagaagaagatgctgtaggaagatgctcctaagca 1207
DB 1647 CTGCGAGTGAAGTGAAGTGAAGGAGAAAGAACCGCGGTGATCACCGTTGACCGCGA 1706
OY 1208 caagagctgaggttgaagaaggtatgctccctgagtgagtgagctcttgtagcgcctca 1267
DB 1707 CTAAGCGCGCGTGAAGAGCGATTTGATTTGGGCGGTGCGGCGCTCATTCGCGCG 1766
OY 1268 ttaagctctgtagatataaactcgtgtagatgtagaactgctgtaactatca 1327
DB 1767 CCGCAAAAGTGCATTTGATTTTAC-----ACGATGATGAAGAAAGTGGGCTATGAATCA 1820
OY 1328 tccgtgctctcttgaagaagccttaagctcaaatgctgcaaatgctgtagtaagct 1387
DB 1821 TCATGCGCGCATTTAAACCCCATTAAGCTCAAAATGCTATCATCCGTTATGATGCG 1880
OY 1388 ctatgtgtgtagaagaatctgtagaacaagaatggtttgtagttaaagctgtagcag 1447
DB 1881 GTGTGCTGTGAATGATGATGAGAAACAGAGGCGCATTTTGTGTTAAAGCTTACCAATG 1940
OY 1448 gagaatctgaagaacttaagaagctgtagtcaactcctaagaagaactcaagctatg 1507
DB 1941 GCAATATGCTGACATGTTTAAAGAGGATATTGACCCCTTAAATGATGAAGAGATCG 2000
OY 1508 catcaaaaatgtagcagcatcagtagccctcctactccttaactcaactagaatgtagctatg 1567
DB 2001 CTTTACAAAATGCGGTTTTCGTTTCAAGCTCTGCTTTTAAACCAAGAGCCACGTCATG 2060
OY 1568 aaaaaccgaacttaaaaagataatgctatgctg----gagtgtagtggtagtggtag 1624
DB 2061 AAATCAAAAGAAAGAAAGCGCGCCAGCAATGCTGATATGAGTGGCATGCGCGGAATCG 2120
OY 1625 gtagtagcagtagtacta 1646
DB 2121 GAGGCTAGGGCGCATGATGTA 2142
```

RESULT 7
US-08-432-697-28
Sequence 28, Application US/08432697
Patent No. 6248330
GENERAL INFORMATION:
APPLICANT: Labigne, Agnes
APPLICANT: Sauerebaum, Sebastien
APPLICANT: Ferreiro, Richard L.
APPLICANT: Thibierge, Jean-Michel
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farbow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/432.697
FILING DATE: 02-MAY-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495.0137-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ. ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 2284 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-432-697-28

Very Match 37.1%; Score 610.4; DB 4; Length 2284;
est Local Similarity 62.1%; Pred. No. 5,2e-152;
Matches 1019; Conservative 0; Mismatches 611; Indels 12; Gaps 3;

QY 8 cttaagaatcccttttgatgctaaagcccgtagaaaccttccagagtgtagataac 67
DB 510 CAAAGAAATCAATTTTCAGATAGCCCAAGAAACCTTTATTGAGCGGTAAAGCAAC 569
QY 68 ttgcaaatgctgttaagaatgacacttgacctaaggcgtlaagtcgtattgaagaag 127
DB 570 TCCATGACCGCTGTCAAAGTAACTGGGGCCAAAGCCAGCAAGCGTTTGAATCCAAAA 629
QY 128 ctttggctcccaagttatcaaaagatggtglatcgtgtgcaaaagaaattgaactg 187
DB 630 GCTATGGCGGCTCCAAAGCATCAACCAAGCGCGTGAAGCGGTAAAGAGATTGAATTA 689
QY 188 aagataagtttgaagaatggtggtcctaagaatggttgaagaatgagtcctcaaaactag 247
DB 690 GTTGGCCCGTGAATGATGGCGCTCAAGCTGTTAAAGAGATGCAAGCAAAACCCCTG 749
QY 248 atattgctgtgtagaagacacacagcaacagtccttgacaaagctattatcgtgaag 307
DB 750 ATGCCCGCGCATGACACACACACACACCGCTGCTGCTTATAGCATTTTAAAGAGG 809
QY 308 ggttaaaacttgtagcagctggtgctgaatcctlaagccattaaagctgacatagataag 367
DB 810 GCTTGAAGATATCAGCGCTGGGGCTAACCCTTATTGAAGTGAAGCAAGCGCATGATTAAG 869
QY 368 ctgttgtgtgttactaaagaactaagcgacattacaagaagcctactcgtgaccaaag 427
DB 870 CGCGCTGAAGCATATTAAAGACCTTAAAGACGACGCAAAAGATGGGGCGTAAAGAG 929
QY 428 aatagtcacagcttgagacacattctgcaaaccttgatcaacaatagatgaatcctag 487
DB 930 AATATCAACCAAGTACGACCATTTTCTGCAACTCCGATCAATATGCGGAACCTCTGCG 989
QY 488 ctgaagctatgctaaagttgaaagaagagtggtatcacagctgaaagaaactaaagtc 547
DB 990 CTGACGCTATGGAAGAAAGTGGTAAAGAGCGGCTGATGACGTTGAAGAGGCTTAAGGCA 1049
QY 548 ttgaactacatagatggtgtggaagaatgaaattgacggtgctactcctccat 607
DB 1050 TTGAAGATGAATTGATGCTGTAAGAGCGATGCAATTGTAGAGGCTACCTCCCTT 1109
QY 608 acttgaactaaactgaaagaatggttgtagaactgaaacccctatcattcgtga 667
DB 1110 ACTTTGTAACCAACCGCTGAAAGAAATGACCCGCTCAATTGATTAACGCTTTTAA 1169
QY 668 atgagaagaagatlaactagcaagaagacatgctaccaatctagaacaagtgctaaag 727

DB 1170 CGGATAAAAAATCTAGCATGAAGACATTTCTCCGCTACTGTGAAAAAACCTGTAAG 1229
QY 728 taacgcgccacccctctattatctgtaagaacgtagaagtgaaacacttgcaacactg 787
DB 1230 AGGCAAAACCGCTTTTATCATCTCGTGAAGACATTGAAGCGGAAAGCTTTAAAGCATCTAG 1289
QY 788 tagtcaatlaagctccggtggaacactccaagcttgtagccgtaaaagctcctggttgcgt 847
DB 1290 TGTGTAATTAATTAAGAGCGCTGTTGAATATCCAGCGCTTAAAGCTCCAGCGCTTGGCG 1349
QY 848 aacgcgtaaagctatgtaagaatattgctacccctactcgtggaaggaagcaatattg 907
DB 1350 ACAGAGAAAGAAATGCTCAAAAGACATCGCTGTTTAAACCGCGCTCAAGTCAATAGCG 1409
QY 908 aagatcgtgtataaagcttgaagaatgtaagctgtcctctttagaagaactaaagctg 967
DB 1410 AAGAAATGGCGCTTGTGCTGTAAGAAACCTGAGTGAAGTGTGTTTAAAGCAAGCGAA---GA 1466
QY 968 tagtlatgacaaagaataactatcgtttagatggtgctggaataacgaagata 1027
DB 1467 TTGTGATTGAACAAAGACACACCGATGATGCAAGGCAATAGCCATGACCTCA 1526
QY 1028 aagctgagttaaacaattcgtgcacaattggaagaagaagctcagatcagatgctg 1087
DB 1527 AAGACAGAGTCGCGCAATCAAAACCCAAATTCGACACGACAGCAAGGATTTACCAAAAG 1586
QY 1088 aaaaacttcaagaacgcttctgcaaaacttggttgtagtgaagctgtatccatgttgag 1147
DB 1587 AAAAATGCAAGAAAGATTTGGCCAAACTCTGTGCGGTGTCGCTGATTAATAGTGGCG 1646
QY 1148 ctgtactgtaactgaaatgaagaagaagaagatcggtgtagaagatgctcctaaatgcaa 1207
DB 1647 CTGCGATGGAAGTGAATTAAGAAAGAAAGAAAGCCGCGTGAAGTACGCTTGACCCGGA 1706
QY 1208 caagagctgctgtgaagaaggtatctccctggtggtggtgtagcgtcttgctgcacca 1267
DB 1707 CTAAAGCGGGGTTGAAGAAAGCATGTGTGTTGGGGCGGTGCGCCCTATTGCGCGCG 1766
QY 1268 ctgaagctccttgatgatatataaactgctgtagatgaagaaactgctggaactaataca 1327
DB 1767 CCCAAAGAGTGCATTTGAAATTTTAC-----ACGATGATGAAGAAAGTGGCTATGAATCA 1820
QY 1328 tccgctgctccttgtagaagccttccctgtaaatctgtaaatctgtaaatctgtaaatctg 1387
DB 1821 TCAATGCGCGCATTAAGACCCCATTTAGCTCAATTCGCTATCAATGCGGTTATGATGCGC 1880
QY 1388 ctatgtctgtagaagaagctgtaacccaagaagatggtttagatgaatgctgctgcatcag 1447
DB 1881 GTGTGCTGATGAAGTGAAGAAAGCAAGCAAGGCGATTTGGTTTAAAGCGTAAAGCAATG 1940
QY 1448 gagaataagaaacctatataaagctggtgctatgatacctaaagaagttacacgtatg 1507
DB 1941 GCAAGTATGTCGATGATTTAAAGAAAGGCAATTTATGACCCCTTAAAGATGGAAGAGT 2000
QY 1508 cattacaagaatgtagacatgtagtcctccttacttcttaactaagaatgtagcattcgtg 1567
DB 2001 CTTTACAAATGCGGTTTCCGTTTCAAGCCTGCTTTAAACCAAGAGCCACGCTGCATG 2060
QY 1568 aaaaacgaacccaataaagaatgctcctgctg---gcggtgtagtgggtggtatg 1624
DB 2061 AATATCAAGAAAGAAAGGCGCCCAAGCAATGCTATATGATGATGCGATGCGGGAATG 2120
QY 1625 gtagtatggaaggtatgacta 1646
DB 2121 GAGGCAATGGCGCGCATGATGTA 2142

RESULT 8
US-08-461-775-10
; Sequence 10, Application US/08461775
; Patent No. 5858773
; GENERAL INFORMATION:
; APPLICANT: MAZODIER, Philippe

```

APPLICANT: GUGLIEMI, Gerard
TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE
TITLE OF INVENTION: INITIATION OF TRANSCRIPTION
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461.775
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/050.313
FILING DATE: 10-MAY-1993
APPLICATION NUMBER: FR 9011186
FILING DATE: 10-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 010830-035
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1620 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1620
US-08-461-775-10

```

```

very Match 24.1%; Score 396.6; DB 2; Length 1620;
est Local Similarity 54.3%; Pred. No. 1.5e-95;
Matches 848; Conservative 0; Mismatches 704; Indels 9; Gaps 2;

```

```

OY 22 tttagatgtaagccgctgaaataaacttcacgaggtgtagatgaacttgcgaatgctgtt 81
DB 19 tttagatgtaagccgctgaaataaacttcacgaggtgtagatgaacttgcgaatgctgtt 81
OY 82 aaagttaacttgcgaatgaaagccgtaattgctgtaattgaataaagcttcttggcccca 141
DB 79 AAGGTGACCAATGCGCCGAGGCGGCAAGCGTCGTCATGCAACAAGATTGGGCGCCCG 138
OY 142 gtataataaagaatggtatcgtctgcaaaagaataatgaacttgaagataaagttgaa 201
DB 139 ACCATCCCAACGACGAGGCGGTACCATCCCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 198
OY 202 aataatgagcgtcgaatggttaagaagtagctcccaaaactgaagataatgctgtgagt 261
DB 199 AACCTCGGCGCCACCTGTCAGAGAGGTGGCGACCAAGCAACATCGCGGTGAC 258
OY 262 ggaactacaagacaagctccttcgacaaagtattatcgttgaagtgtaaaacttgta 321
DB 259 GGCACACACACCGACCGGTGTGCGCCAGAGGCGTGTGCGCGAGAGGCGCTGGCAAGTC 318
OY 332 gcaagctggtcgttaactcctatgagcaataaagctgagataaagctgctgtgctgt 381
DB 319 GCCCGCGCGCTCCCGCGCGCTGTGAAGAGGCAATGACGCGCGCTGCGCGCTG 378

```

```

OY 382 actaagaactaagcgacattacaagaagcctcgttgacaaagaagaatgactcaagt 441
DB 379 TCCGCGGACCTCTCGACACCGCGCGCGGATCGACGACAAATCCGACACTCCCGCGCTC 438
OY 442 ggaaccattctgcgaactcgtatacaacaagaatgaatataatagctgaagctatgct 501
DB 439 GCCGCGCTCTCCCGCGAG---GACAAGCAGTGTGGCGAGCATGTCGCCGAGCGATGAC 495
OY 502 aaagtgtgaagaagggtgttatccagcttgaggaaagctcaaaagcttgaataactcata 561
DB 496 AAGGTGCGCAAGGACGAGGTTCATCACCGTCGAGAGGTCCAAACCTTCGTTGCAACCTG 555
OY 562 gactggttgaagaagaaagcttgcacgtgagcctcctcccaacttgaacttaact 621
DB 556 GACTTTCACCGAGGCGATGCGCTTCGACAGAGGCTACTCTCCCGTACACTGGTGACCGAC 615
OY 622 cctgaagaatggttctgtaactcgtataacccctatcctcttgtaattgaagaagaat 681
DB 616 CAGGAGCGATGAGAGCGCTCCTCGACGACCGGTACATCTGATCCACGAGGCAAGATC 675
OY 682 actagatgaagaagacatgctaccaatctagaacaagt-----tgcataagtaacgt 735
DB 676 GGTTCGATCCAGGACCTGCTGCGCTGTGAGAGGTATCCAGGCGGGTGTCCAG 735
OY 736 ccaactctatattatgctgaaagcgtagaagtgaaagcacttgcacacttgaactcaat 795
DB 736 CCGTCTGATCATATGCGCGGAGACGTCGAGGCGGAGGCGCTTCGACCTGGTGTCAAC 795
OY 796 aagctcgttgaagcactccaaagtgtgacgtaaaagctcctggttctgttgaagcgt 855
DB 796 AAGATCCGCGGACGTTCAAGCGCGTCCGCGTCAAGGCGCGCGGCTGTGTTGATCCGCGC 855
OY 856 aaagctatgcttgaagaatatactccttactcgtgaaggaagaacataattgaagatcgt 915
DB 856 AAGGGATGCTCGGCGACATGCGCACCTCACCGGTGCCACCGTATCCCGACGAGAGTTC 915
OY 916 gataaagcttgaagaatgtaagctgtctctcttgaagaagcctaaagctgtaagt 975
DB 916 GGCCTCAAGCTTGACCGAGCGCGGTGTGACGCTGTGGACGCGCGCGCGCTGTCAACGTC 975
OY 976 gacaagaagaataactactcctcgttgaagtgctggaagaataactgaagaatgctga 1035
DB 976 ACCAAGGACGACGACGACCATCTGAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1035
OY 1036 gttaacaataatcgtgcacaaattgaagaacaagctcaagattatgactgtaaaaaact 1095
DB 1036 GTCCGCCAGATCAAGCGGAGATCGAGTCGACGCACTCGACTGGGACCGCAGAGAATTC 1095
OY 1096 caagaagctcttgcgaacttctgttgagtagtctgtatcactatgtagtgaacttact 1155
DB 1096 CAGGAGCGCTGCGCAAGCTGCGCGCGCGCGCTGCTGCTGATCCGCGTCCGCGCGCAC 1155
OY 1156 gaaactgaatgaagaagaagaagatcgttgaagaagatgctcctaattgcaacaagagct 1215
DB 1156 GAGGTGAGGTGAAGAGGCGGACGACCGCTGTGAGAGGACCATCTCCGCGACCGCGCGC 1215
OY 1216 gcggttgaagaagatgctcctggttgaagctgcttctgtcgtcgtcattaaagtc 1275
DB 1216 GCGGTGAGAGGAGGCGATGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1275
OY 1276 cttagatlatlaaacccgtgtatgatgaactgtcgtgacttaataataatccgttgt 1335
DB 1276 CTGGACGACAACTCTGCGCGCGACCGCGCGACGAGCGCACCGGCTGCTGCTGCTGCTGCT 1335
OY 1336 tctcttgaagagccttactcgttaaatctgctgaagaatgctgctgaaggtctatgct 1395
DB 1336 GCGCGCGTGAAGCGCGCTGCTGAGTCCGCGAGAAAGCGCGCTGAGAGGCTACGTCAATC 1395
OY 1396 gtaagaagaatgctgtaacaaagaatggttctgtgaattatgctgctcagagaagat 1455
DB 1396 ACCACCAAGGTGGCGGACCTGACCAAGGCGCGGCTTCAACGCGCGCACCGCGCAATAC 1455
OY 1456 gaagacctatlaaagcgtgtgctatgctccttaaaaaagtaacgctatgctatcaaa 1515

```

Db 1456 GGGGACCTGTCAGGCGCGGTCATGACGCCGTCACAGCTACCGCCCTGGAG 1515
Qy 1516 aatgcagcaatcagtagctcttacttaactaagaatgcatatgtgaaaaacca 1575
Db 1516 AACGGCGGCTCCATGCGCTCCCTGCTGACGACGAGACCCCTGCTCGAAGAGCGG 1575
Qy 1576 G 1576
Db 1576 G 1576

RESULT 9

US-09-031-606-10
Sequence 10, Application US/09031606
Patent No. 6153404
GENERAL INFORMATION:
APPLICANT: MAZODIER, Philippe
APPLICANT: GUGLIEMI, Gerard
TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031.606
CLASSIFICATION:
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/050.313
FILING DATE: 10-MAY-1993
APPLICATION NUMBER: FR 9011186
FILING DATE: 10-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 010830-035
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1620 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1620
US-09-031-606-10

Query Match 24.18; Score 396.6; DB 3; Length 1620;
Best Local Similarity 54.38; Pred. No. 1.5e-95;
Matches 848; Conservative 0; Mismatches 704; Indels 9; Gaps 2;

Qy 22 ttatattcaaaagccgtgaaaacattcaggaagtgtagataaactgtgcaatgctgt 81
Db 19 TTCACACAGACAGCCCGCTGCGCCCTTGAGCGCGCGGTGAACCACTGGCCGACACCGTC 78
Qy 82 aaagtaaacacttgagacctaagaagcgtaatgtcgtattgaaaagctcttgggtccca 141

Db 79 AAGGTGACCATGTCGGCCCAAGGCGGCACAGTCGTGATCGACAGAGATTTCGGCCCGCC 138
Qy 142 gtattcaaaaagatggtgtatctgttgaaaaagaatgtaactggaataatgttga 201
Db 139 ACCATTCACCAAGACAGCGGCTCACATCGCCCTGAGAGTGAATGGACACACCGGTAAGC 198
Qy 202 aatatggcgctcaaatgtttaaagaagttagctcccaaaactagcatalgtgtgtat 261
Db 199 AACCTGGGCGCCACACTGCTCAAGAGAGTGGCGACCAAGACACACACTCCGGGTGAC 258
Qy 262 ggaactacaagcagcaagctcttgacaagctattatcgtgaagtgtgaaactgtga 321
Db 259 GGACACACACCGCACCGTGTGTCGGCCAGGCGTGTCCGGAGAGGCGCTCGCAACCTC 318
Qy 322 ggaagtgtgtgaatcctatcctatgcaatgaagctgtgcatagaataagctgtgtct 381
Db 319 GCGCGCGCGCGCTCCCGCGCGCGCTGTAAGAAAGGCAATGAGCGCGCGCTCCCGCTC 378
Qy 382 actaaagaactaagcagcatatcaaaagcctactctgacaaagaataatgctgaagt 441
Db 379 TCCGCGGAGCTGCTGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 438
Qy 442 ggaacatattctgcaaacctctgatacaacaataagtaataatcagctgaagtatgct 501
Db 439 GCGCGCTCTCCGCGCAG--GACAAAGCAGGTGCGCGCGCTCATCGCGAGCGCATGAGC 495
Qy 502 aaagtgtgaaaagaaggtgtatcacaagctgagaagctgaagctgtgaactacata 561
Db 496 AAGTCGCGCAAGAGAGCGTGTATCATCCCTCGAGAGTGTCAACACCTTGCTGTCACTG 555
Qy 552 gatgtgttgaagaagaatgttgaaccgtgtactctctcacttctgataactatgta 621
Db 556 GACTTCAACGAGGCGCATGGCTTCGACAAAGGCTACCTGTCGCCGTACATGCTGACGAC 615
Qy 622 cctgagaagaatggttgaacttgataacctatattccttgaatgagaagaagaatc 681
Db 616 CAGGAGCGTATGAGAGCGCGCTCTGACGACCGGTATCATGATCCACAGGCAAGTAC 675
Qy 682 actagcatgaagaacatgtctacccaactctagaacaagt-----tgctaaagtaaacgt 735
Db 676 GGTTCGATCCAGACACTGCTCCCTGCTGAGAGAGTCAATCCAGCGCGGCTGCTCCAG 735
Qy 736 ccactccttatttctgtgaagcgttagaaggtgaagcacttgcaactgtgataact 795
Db 736 CCGCTGCTGATCATCGCGAGAGAGTGAAGGCGCGCGCGCGCGCGCGCGCGCGCGCT 795
Qy 796 aagctcgtgtgagaccccaagctgttagcgttaaaagctcctgtgttgtgtgaagcgt 855
Db 796 AAGATCCGCGGACGTTCAAGCGCGTGCCTCAAGCGCGCGCGCTTCGGTGACCGCGC 855
Qy 856 aaagctatgtctgaagatattgtctactccttactgaagagaagaataattgaaagct 915
Db 856 AAGCGATGCTCGCGCGGACATGCGCACTCCCTGCTGACCGCTGCTGCTGCTGCTGCT 915
Qy 916 ggtataaagcttgaaaatgtgaagctgtctcttgaagaaagcgtgaagctgtgatat 975
Db 916 GCGCTCAAGCTGACACGAGCGCGCTGTGAGAGCTGCTGCGGACCGCGCGCGGTGACCGTC 975
Qy 976 gacaagaaaaataactactatcgtgtgagtggtgtgaaagacagaatattaaagctga 1035
Db 976 ACCAAGAGACAGACGACCATGTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1035
Qy 1036 gttaaacaatctgtgacaaatgtgaagaacaagctcagataatgatactgtgaaaaact 1095
Db 1036 GTCCGCCAGATCAAGCGCGGAGATGAGATCGACCGACACTCGGATCGGACCGCGAAGACTC 1095
Qy 1096 caagaacgtcttgcacaaactgtgtgtgagtagcgttatacaltgttgaagctgtact 1155
Db 1096 CAGAGCGCTTCGCGCAAGCTGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1155
Qy 1156 gaacatgaatgaagaagaagatcgtgtgagaatgcttaaatgtaaaagaagct 1215
Db 1156 GAGGTGAGCTGAAAGAGCGCAAGCACCGTCTGAGAGACCATCTTCGCGCGCGCGCGC 1215

Yy 1216 gcggtlgaagaagatctccctggtagtgtaacttctttgctcgcgtcatlaaaagtc 1275
|| || || || || || || || || || || || || || || || || || || || || ||
Db 1216 GCGGTCAGACGAGGCCATCTCTCCGGTGTGGCTTCGCCGCTTGCTCACGCCCTCAAGGTC 1275

Qy 1276 ctgatgatataaaccctgcctgatgatgatgaacttgcctggaactaatatcatccgctgt 1335
|| || || || || || || || || || || || || || || || || || || || || ||
Db 1276 CTGAGACGCACACTCTGGCGCGCACC GGCGACGAGGCCACCGGATGTTGGCGGTGTCGGCGCGC 1335

Oy 1336 tctcttgaaggagccttcaatgcctcaaatgcctcaaatgcctcaatgcctcaatgcctcaatgc 1395
|| || || || || || || || || || || || || || || || || || || || || ||
Db 1336 GCCCGCCGTCGAGCGCGCTGCGCTGGATCGCCGAGAACGCGCGGCTCGAAGGGCTACTCATTC 1395

Yb 1396 gtgaaaaagtctgtgaaccaaaaagtgtgttttgatttaatgtcgtacacaggaagaatc 1455
|| || || || || || || || || || || || || || || || || || || || || ||
Db 1396 ACCACCAAGAGTGTGGCGGAECTGCACAGAAGCCAGAGGCTTTCACAGCGCGACCGGCACAGTAC 1455

Jy 1456 gaagaccttatlaaagctgtgtgcatltgatcctaataaaagtacacgatatgtatlaaca 1515
|| || || || || || || || || || || || || || || || || || || || || ||
Db 1456 GGGCAGACCTGTCAAGAGCGCGGGTATCGACCCGGTCAAGGTACACCGCGCTCCGCCCTGGAG 1515

Oy 1516 aaatgcagatcagtagagctcctcttaacttaactaacagaaatgcgtatgtcgtgaanaaca 1575
|| || || || || || || || || || || || || || || || || || || || || ||
Db 1516 AACCGCGCTCATCTGCTGCTCCCTGCTCTGACGACGACCGACCTGTGTGTGAGAACCG 1575

Qy 1576 g 1576
|
Db 1576 C 1576

RESULT 10
US-08-997-080-113
Sequence 113 Application US/08997080
Patent No. 5968524
GENERAL INFORMATION:
APPLICANT: WATSON, JAMES D.
APPLICANT: TAN, PAUL L.J.
TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY
NUMBER OF SEQUENCES: 194
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,080
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sleeth, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1007
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 1569 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA

| Query Match | 23.7% | Score 390.8 | DB 2 | Length 1569 |
|-----------------------|---|-------------------|----------|-------------|
| Best Local Similarity | 53.9% | Pred. No. 4.9e-94 | | |
| Matches 847 | Conservative 0 | Mismatches 717 | Indels 6 | Gaps 2 |
| QY 8 | ctlaaagaalcccttttgatgctaaagcccgctlgaaaaacttcacgaggctgataaac | 67 | | |
| DB 5 | CCAACAACAATTGCGTATGACGAAAGAGCCCGCCGCTGCATCGACGGGGCCCAACGCC | 64 | | |
| QY 68 | ttgcaaaagctggttaagtaaacacttggaccctaaaggccgctaatgttgcttatgaaagt | 127 | | |
| DB 65 | TCGCAAGCGCGCTAAAGGTGACGTGGGCCCCCAAGGGTGGCAACCTCGTGGTGGAGA | 124 | | |
| QY 128 | cttttggtcccgagttacttaacaaaagtgtgtatctcgttgcataaagaatctgaactg | 187 | | |
| DB 125 | AGTGGGGCGGCCCCCAACGATACACACGATGCTGTCCATCGCCAAAGAGATCGAGCTGG | 184 | | |
| QY 188 | aagataagtttgaataatatyggcgctcaaatlygtttaaagaagtagctcccaaatagcg | 247 | | |
| DB 185 | AGGACCCCGTACGAGAAGATCGCGCGCTGAGCTGTCAAGAGAGTCCCAAGAAAGCCGACG | 244 | | |
| QY 248 | atatgtcgtgtatgtgaactatacaagaacagctccctgcacaaagtattatgtgaag | 307 | | |
| DB 245 | ACGTGGCGGGCCACGGCACACACACCGCCACCGTCTGCTCAAGCTGTGGTCCCGAAG | 304 | | |
| QY 308 | gtgtaaaactgttagcagctgtgtgttaactcctatgacctaaaggttgacatagaag | 367 | | |
| DB 305 | GCCTCGCAACGCTCCGACCGGGCCCAACCGGCTTGCCCTCAAGGCTGGCATCGGAAGG | 364 | | |
| QY 368 | ctgtgtgtgtgttacttaaaagaaactaagcgaaattacaagagctactcgttgaccaaaag | 427 | | |
| DB 365 | CTGTGACGAGCTGTCAACCCAGTCCGTGCTGTAAGTGGCCCAAGAGGTCAAGCAAGAGAC | 424 | | |
| QY 428 | aaatagctcaagcttggaaacattctctgcaaatctgtatacaaaatagtgtaataatag | 487 | | |
| DB 425 | AGATTTCAGCCACCGCGCGATTCGCCGCGCG --- GACACCACGATCGCGAAGCTCATCG | 481 | | |
| QY 488 | ctgaagctatgtgctaaagtltgaaaaagaagtgttatcaacglttgaaggaagctaaagtc | 547 | | |
| DB 482 | CCGAGGCCATGACAAAGGTGGCAACGAGGGTGTATCACTCCGTCAAGGAGTCTGAAACACT | 541 | | |
| QY 548 | ttgaacctacatlaagatgtgttgaagaaatgaaglttgaacggttgacactctccat | 607 | | |
| DB 542 | TGCGGCTTCGACGTCGACGCTCACCGAGGTTATCGCTTGACAAAGGCTACATCTGGGTT | 601 | | |
| QY 608 | acttltgtaactaatccgtgagaanaatlygtlttgbaaacttgataaaccttatatccttltg | 667 | | |
| DB 602 | ACTTGTGACGACGACGCCGACGCGCAGGAAGCGTCTGTGAGATCCTACATCTCGTGG | 661 | | |
| QY 668 | atgagaaaaaagaattctctgcatgataagaagcatgtctacaaactttagaaacagtgtctaag | 727 | | |
| DB 662 | TCAGCTCCAAAGGTGTGACCGGTCAAGGATCTGCTCCCTGCTGGAGAAAGTCAATCCAGG | 721 | | |
| QY 728 | taaacgctccactctcttaatttgttgaagacgtataagaagtgaaagcacttgcacaacttg | 787 | | |
| DB 722 | CCGGCAAGCCGCTGCTGATCATCGCCGAGAGAGCTGAGAGGCCAGGCCCTGTCCACGCTGG | 781 | | |
| QY 788 | taagtcaataaagctccgttgaagcactccaagtltgtaacggtataaagcttcctgttltgtg | 847 | | |
| DB 782 | TGCTCAACAAAGATCCGGGACACTTCAAGTCCGTGCGGTCAAGGCTCCGGGCTTCGGTG | 841 | | |
| QY 848 | aacgcctgaagaactgtgcttgaagataatgtctacacctactcgtgaaggaagaacatatctg | 907 | | |
| DB 842 | ACCGCGGAGAGCGATGTGTGACGACATGTGGCCATCTCAACGGTGTGATAGTCTGCACGG | 901 | | |
| QY 908 | aagatcctgtgtataaagcttgaanaatgttaagctgttctctttaggaacacgttaaagtg | 967 | | |
| DB 902 | AAAGAGTGGGGTGTCTCTGGAGACGCCCGAGAGTCTCCGTGGGCCAAGGCCCAAGC | 961 | | |
| QY 968 | tagttattgcaaaaataataactactactgttgaatgtgtgttgaanaatcagaagataat | 1027 | | |

```

Db 962 TCGTCCACCAAGCAGACCAACATCTGAGGCGCTGGCGATTCCGATGCCATCG 1021
Qy 1028 aagctcaggttaaacaaatltgtgacaaatltgaagaacagctcaatattatgtg 1087
Db 1022 CCGGCGGGGTGCTAGATCCGCGCGGATGAGAACAGGATCCAGTACGACCCGCG 1081
Qy 1088 aaaaactcaagaacgtcttcaaaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1147
Db 1082 AGAAGCTCAGAGAGCGCTGCGCAAGCTGCGCGCGGTGTGTGTGTGTGTGTGTGTGT 1141
Qy 1148 ctgtacttgaactgaaatgaagaagaagatcgtgtgtgtgtgtgtgtgtgtgtgtgt 1207
Db 1142 CTGCCACCGAGCTGAGCTCAAGAGCGCCAGACCCGATGAGAGAGCGCGCTGCCAGC 1201
Qy 1208 caaagagctgaggttgaagaagaatltgtccctgtgtgtgtgtgtgtgtgtgtgtgt 1267
Db 1202 CGAAGGCGCGCTGAGAGAGGAGGATGCGCGCGGTGTGTGTGTGTGTGTGTGTGTGT 1261
Qy 1268 ttaaagctctgtatgtatataaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1327
Db 1262 CTCTGCGCGCTGAGAGAGCTGCGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1318
Qy 1328 tccgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1387
Db 1319 TCCCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCG 1378
Qy 1388 ctatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1447
Db 1379 GCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1438
Qy 1448 gagaataatgaagaacttataaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1507
Db 1439 GTGAGTGAAGAGAGCTGCTCAAGAGCGCGCGCTGAGAGAGAGAGAGAGAGAGAGAG 1498
Qy 1508 catcaaaatgaagaatcagtagctccttacttacttacttacttacttacttacttact 1567
Db 1499 CCGTCAAGAGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1558
Qy 1568 aaaaaccaga 1577
Db 1559 ACAAGCCGGA 1568

```

```

RESULT 11
US-08-997-362-113
: Sequence 113, Application US/08997362
: Patent No. 5985287
: GENERAL INFORMATION:
: APPLICANT: Tan, Paul
: APPLICANT: Hiyama, Jun
: APPLICANT: Visser, Elizabeth
: APPLICANT: Skinner, Margot
: APPLICANT: Scott, Linda
: TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
: TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
: NUMBER OF SEQUENCES: 194
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Law Offices of Ann W. Speckman
: STREET: 2601 Elliott Avenue, Suite 4185
: CITY: Seattle
: STATE: WA
: COUNTRY: USA
: ZIP: 98121
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/997,362
: FILING DATE:
: CLASSIFICATION:

```

```

PRIOR APPLICATION DATA:
: APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
: FILING DATE: June 12, 1997
: APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
: FILING DATE: August 29, 1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Sleath, Janet
: REGISTRATION NUMBER: 37,007
: REFERENCE/DOCKET NUMBER: 11000,1002c2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 206-269-0565
: TELEFAX: 206-269-0563
: TELEX:
: INFORMATION FOR SEQ ID NO: 113:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1569 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: Genomic DNA
: US-08-997-362-113

```

Query Match 23.7%; Score 390.8; DB 2; Length 1569;

Best Local Similarity 53.9%; Pred. No. 4.9e-94; Matches 847; Conservative 0; Mismatches 717; Indels 6; Gaps 2;

```

Qy 8 cttaaagaatccttlttgaatgtttaaagccgttgaagaacttcaagaagtgtatgaataac 67
Db 5 CCAAGACAAATGCGTATGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 64
Qy 68 ttgaagaatcgtttaaagaatgaacttgaacttgaagaagcgtgaatgtgtatgtgaagaat 127
Db 65 TCGGAGAGCGCGTAAAGGTGACGTTGGCCGGAAGGGTCCGCAAGCTGCTGTGAGAGAGA 124
Qy 128 ctttgttcccgagttattcaagaagatgtgtatcgtgtgtgtgtgtgtgtgtgtgtgtgt 187
Db 125 AGTGGGGCGCCCGCCAGCATACCAAGATGATGATGATGATGATGATGATGATGATGATGAT 184
Qy 188 aagaataatgttaaaatattggtgtcaaatgtttaaagaagtagctcccaaaactagcg 247
Db 185 AGGACCCGTACGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 244
Qy 248 atattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 307
Db 245 ACGTGGGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 304
Qy 308 gtgtaaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 367
Db 305 GCGTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 364
Qy 368 ctgttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 427
Db 365 CTGTGAGAGCTGTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 424
Qy 428 aaaaagctcaagtgtgaaccatttctgtcaaatctgtatcaacaataagtgtaatacatag 487
Db 425 AGATTGTGCGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
Qy 488 ctgaagctatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 547
Db 482 CCGAGGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 541
Qy 548 ttgaactatagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 607
Db 542 TCGGCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601
Qy 608 acttgaactaatcctgtgaagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 667
Db 602 ACTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 661
Qy 668 atgaagaagaattactagatgaagaagacatgtaccactttagaacaagtgtgtgaag 727

```

Db 662 TACCTCCAGCTGTGACCTGCAAGATGTGCTCCCTGCTGAGAGGTATCCAGG 721
Qy 728 laaacgclcaactctctatlatlgtcgaagacgtagaagtgtaaaccttcaacttg 787
Db 722 CCGGCAAGCCGCTGTATCTATCTCCGCAAGAGCTGAGAGGAGGCCCTGTCCAGCTGG 781
Qy 788 tagtcaataagctcgtgtagcactccaagtgttagccgtaaaagctcgtgtttggtg 847
Db 782 TGTGTACACAGATCCGGGACCTTCAAGTCCGTCCGTCAGGCTCCGGGCTTCCGTTG 841
Qy 848 aagcgtaaaagcgtacgttgaagatattgtctactcctcgtgagaggaagaacataattg 907
Db 842 ACCCCCAAGCAGCTCTGACAGACATGCGCATCTCACCGGTCGTACGTCTCAACCG 901
Qy 908 aagctcgtgtatataagcttgaagatgttaagctgttcttctttaggaacgtataagtg 967
Db 902 AAAGAGTGGGCTGTCTCTTGAGACCCGCCAGCTCTCTCTGAGGCCAGCCCGCAAGG 961
Qy 968 tagtattgaagaagaataatactatactgttgaatgtgtgtcgtgaanaatacagaataata 1027
Db 962 TGTGTGTACACAGAGAGAGACACATCTGTGAGAGGCTCGGGGATTCGATGCCATCG 1021
Qy 1028 aagctcgaagtaaaacaaatcgtgcacaaatgtgaagaagaagaactcagatattgctgtg 1087
Db 1022 CCGGCGGCTGCTCAATCCGCCCGGAGATGAGAACAGCCGACTCCGACTACGACCCCG 1081
Qy 1088 aaaaacttcaagaagacgtcttgaagaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1147
Db 1082 AGAAGCTGCAAG 1141
Qy 1148 ctgtctactgaagaactgaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 1207
Db 1142 CTGCGCACGAGAGGTGAGAGGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1201
Qy 1208 caaagactcgtgtgtgaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 1267
Db 1202 CGAAGCGTCCGCTGGAAG 1261
Qy 1268 ttaagactcgtgtgtgaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 1327
Db 1262 CTCTGCGGCTGTGAG 1318
Qy 1328 tccgtcgtctctctgaagaagacgttgaagaagaagaagaagaagaagaagaagaagaaga 1387
Db 1319 TCCGCGTGTGCTGTGAG 1378
Qy 1388 ctattgtgtgaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 1447
Db 1379 CCGTCTGTGCGGAG 1438
Qy 1448 gagaataagaagaacttataaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1507
Db 1439 GTGAGTACAG 1498
Qy 1508 cattaagaataatgcagcatcagtagcctcctacttacttaacacagaagaagcgtatgtg 1567
Db 1499 CGGTGCAAG 1558
Qy 1568 aaaaacaga 1577
Db 1559 ACAAGCCCGA 1568

RESULT 12
US-09-095-855-113
: Sequence 113, Application US/09095855
: Patent No. 6160093
: GENERAL INFORMATION:
: APPLICANT: Tan, Paul
: APPLICANT: Visser, Elizabeth
: APPLICANT: Skinner, Margot
: APPLICANT: Prestidge, Ross
: TITLE OF INVENTION: Compounds and Methods for

1 TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections
2 NUMBER OF SEQUENCES: 208
3 CORRESPONDENCE ADDRESS:
4 ADDRESSEE: Law Offices of Ann W. Spekman
5 STREET: 2601 Elliott Avenue, Suite 4185
6 CITY: Seattle
7 STATE: WA
8 COUNTRY: USA
9 ZIP: 98121
10 COMPUTER READABLE FORM:
11 MEDIUM TYPE: Diskette
12 COMPUTER: IBM Compatible
13 OPERATING SYSTEM: DOS
14 SOFTWARE: FastSeq for Windows Version 2.0
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER: US/09/095,855
17 FILING DATE:
18 CLASSIFICATION:
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: 08/705,347
21 FILING DATE: 29-AUG-1996
22 APPLICATION NUMBER: 08/873,970
23 FILING DATE: 12-JUN-1997
24 APPLICATION NUMBER: 08/997,362
25 FILING DATE: 23-DEC-1997
26 ATTORNEY/AGENT INFORMATION:
27 NAME: Sleath, Janet
28 REGISTRATION NUMBER: 37,007
29 REFERENCE/DOCKET NUMBER: 11000.100263
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: 206-269-0565
32 TELEFAX: 206-269-0563
33 TELEX:
34 INFORMATION FOR SRO ID NO: 113:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 1569 base pairs
37 TYPE: nucleic acid
38 STRANDEDNESS: single
39 TOPOLOGY: linear
40 MOLECULE TYPE: Genomic DNA
41 US-09-095-855-113
42
43 Query Match 23.7%; Score 390.8; DB 4; Length 1569;
44 Best Local Similarity 53.9%; Pred. No. 4.9e-94;
45 Matches 847; Conservative 0; Mismatches 717; Indels 6; Gaps 2;
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

Db 365 CTGTGAGGCTGTACCAAGTCCGTGTAAGTCCGGCCAAGAGGTCAGAGACCAAGAGG 424
Qy 428 aaatagctcaagttggaaccattctctgcaaacctctgatacaacaatagtaatacatag 487
Db 425 AGATTCTGACCAAGCCGGGGATTTCGGCCGGC---GACACCCAGATCGGCGACCTATCG 481
Qy 488 ctgaagctctgctaaagttggaaggaaggtgttaccagttggaaggaagctaaagtc 547
Db 482 CCGAGGCAATGACAGAGTCCGACAGAGGTTCTATCACCCTGAGAGTCCAGACACT 541
Qy 548 ttgaacctacatagatgtgtgttgaaggaagtgttaccgttgacctctccat 607
Db 542 TCGGCTGAGCTGAGCTGACAGGATGCGGCTTGCACAGGGCTCATCTCGGGT 601
Qy 608 acttctgaactaaacctgagaaaagttgttgaactgtataaccttatcccttga 667
Db 602 ACTTCTGACCGACCGCCGCGGACAGAGCCCTCTGAGAGTCCCTCATCTCTCGT 661
Qy 668 atgagaaaagattactagaagaacatgctaccacatctttagaacaagttgctaaag 727
Db 662 TCAGCTCCAGAGTGTGACCGTCAAGATCTGCTCCGCTGCTGAGAGAGTATCAGG 721
Qy 728 taacacctccactcttatattgtctgaagacgtgaaggtgaagacgtgcaacattg 787
Db 722 CCGGCAAGCGGCTGTGATCATCGCGAGAGAGTCAAGGCGCGCTGTCCACGCTGG 781
Qy 788 tagtcaataagctcgtgtgagcaactccaaagtgtgacgtgtaaaagctcctgttgtg 847
Db 782 TGTGTACACAGATCCCGCCACCTCTCAAGTCCCTCCGCTCAAGCTCCGCGCTTCG 841
Qy 848 aacgcgttaagctatgcttgaagatatgcttactgaaggaaggaagaacatlttg 907
Db 842 ACCGCGGCAAGCGGATGCTGCAAGAGATGCGCATCTCACCGGTGTCAGGTCTAGCG 901
Qy 908 aagatctgttgaagcttgaagaaatgtaagctgtctcttcttgaagaaagctaaagctg 967
Db 902 AAGAGCTCGGGCTGTCTGTGAGACCGCGAGCTCTCGGTGCTGGGCGAGCGCCGAGG 961
Qy 968 tagtcttgaacaaagaataactactacatcgttgaaggtgtgtaaaatcagaagatact 1027
Db 962 TCTCTCTACCAAGGACGAGACCATCTCTGAGAGGCTCGGCGCATCTCCGATGCCATCG 1021
Qy 1028 aagctgaagttaaacaattctgtgacaaattgaagaacaagctcagaattatgactcgt 1087
Db 1022 CCGCGCGGCTGTCTGATCTCCGCGAGATCGAGAACAGCGCTCCGACTAGACCGCG 1081
Qy 1088 aaaaactcaagaagctcttgaacaaactgtgtgtgagtagctgtatctctgttgaag 1147
Db 1082 AGAAGCTGAGAGAGCCCTGCGCAAGCTGCGCGGCTGTGCGGTGATCAAGCGCGAG 1141
Qy 1148 ctgtctctgaactgaagaaatggaagaagaaagaaatcgtgtagaagaatgtcttaagca 1207
Db 1142 CTGCCACCGAGGTGAGCTCAAGACCGCAAGCAGCGCATCGAGCGCGCTCCGACAG 1201
Qy 1208 caagagctgaggttgaagaaggtatgtccctgtgtgtgtaactgtcttgcgcctca 1267
Db 1202 CGAAGCTGCGCTGCAAGAGGATGTCTCCGCTGCGGCGGTGCTCTGCTCACTCGG 1261
Qy 1268 ttaaaagcttgaatatactaaacctgtgataagtaagtaactgttgaacttaatact 1327
Db 1262 CTCTCTGCTGAGACGACTGCGGCTGACGGCGAGCAGAG---CCACCGGTCGCAACATCG 1318
Qy 1328 tccgtctgtctcttgaagagctttagcgtcaaatgtctgcaaatgtctgtactagaagtc 1387
Db 1319 TCCGCGTGGCGCTGTCCGCTCCCTCAAGAGATGCTCCCTTCAACGCGCGCTGACCCG 1378
Qy 1388 ctatgttgaagaagaatgtgtaacaaagaatgtgttgaagtaagtaagctgtacatca 1447
Db 1379 GCGTCTGTGCGAAGAGTGTCCACACTGCGCGCGGTGACGCGCTCAAGCGCGGACCG 1438
Qy 1448 gagaataatgaagaaccttataaagctgtgttcatgtatccataaaagaatatacgatgt 1507
Db 1439 GTGAGTACGAGACTGTCTAAGGCGCGGCTCCGACCGCGGTGAAGGTATCACCCTCGG 1498

Qy 1508 catlacaatatgacatcatagtagtaccttacttacttcaactacagaatgctatgctg 1567
Db 1499 CCGTGCAGAACCGCGCTGCATCGCGCTGTCTCTACACCAAGCGCGCTGTGCGCG 1558
Qy 1568 aaaaacaga 1577
Db 1559 ACAAGCCGGA 1568

RESULT 13
US-08-997-080-159
Sequence 159, Application US/08997080
Patent No. 5968524
GENERAL INFORMATION:
APPLICANT: WATSON, JAMES D.
APPLICANT: TAN, PAUL L.J.
TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY
NUMBER OF SEQUENCES: 194
CORRESPONDENCE ADDRESS:
ADDRESSER: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,080
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000,1007.
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 159:
SEQUENCE CHARACTERISTICS:
LENGTH: 1626 base pairs
TYPE: nucleic acid
STRANDEDNESS: Single
TOPOLOGY: linear
US-08-997-080-159

Query Match 23.7%; Score 390.8; DB 2; Length 1626;
Best local Similarity 53.9%; Pred. No. 5e-94;
Matches 847; Conservative 0; Mismatches 717; Indels 6; Gaps 2;

Qy 8 ctataagaatccctttagtctaaagccctgtgaagaaacttcaacgaggtgtatgaataac 67
Db 5 CCAAGCAATTTGGTATGACGAGAGAGCCCGCGCTGCGCTCGAGCGGGGCTCAAGCGCC 64
Qy 68 ttgcaaaagctgtttaaagaacacttggacctaaagcgtaagctgtatgtgaaagt 127
Db 65 TCCGACAGCCGCTTAAGAGTGTACCTGTGCGCGCAAGGCTCGCAAGCTGTCTGAGAGAG 124
Qy 128 ctttggctcccaagttatcaaaagaatgtgtatctgttgaagaaagaatgaagtc 187
Db 125 AGTGGGCGCCCGCCAGATCACCAAGATGtGTGTCTCATCGCAAGAGATGAGCTGG 184
Qy 188 aagataagttgaaaatatggcgctcaaatgtttaaagaagtagtcccaaaactagcg 247

Db 185 AGAGCCCTACGAGAGATCGGCGCTGAGCTGTCAAGAGGTGCGCCACGAGACCGACG 244
Qy 248 aatltgtgtgtatgaaactaacaaagcaagcagtccttgcaagaactatattatgagag 307
Db 245 ACgttgcggcgagcgagcgcacacacacgacgctgtgctcctgacgcttgggttcgagag 304
Qy 308 gtgtataaactgtgacgctgtgtctaaactcctatgacataaactgagcagtagataag 367
Db 305 GCGTGCAGACAGTGCAGCCGCGCCCAACCGCTCGGCTCAGCGTGGCTCGAGAGG 364
Qy 368 ctgttgtgtgttactaaagaactaagcgacattacaaagcctactctgtgacccaaaagt 427
Db 365 CTGTGAGAGCTGTCAACCCAGTGTGAGAGTGGCGCCACAGAGGTGACAGCAAGGAGC 424
Qy 428 aatagtcaagtltgaaacacatltctgaacacttgatacaacaatagtaataatcag 487
Db 425 AGATTTCGCGACCGCGCGGATTTCCGCGCGC---GACACCCAGATCGCCAGCTCATCG 481
Qy 488 ctgaagctatgtctaaagtltgaaagaggtgttatacagatltgaggaagctaaaggtc 547
Db 482 CCGAGGCGATGAGCAAGGTGCGCAACGAGGCTGATCATCCGTGAGAGAGTGAACACACT 541
Qy 548 ttgaaactacattagatgtgtgtgaaagatgaaatttgacgtgtgctactctctccat 607
Db 542 TCGCCCTTCAGCTTCGAGCTCACCGAGCGTATGCCCTTCGACAGGCGTACATTCGCGCT 601
Qy 608 actltgaaactaatctgagagaagaalgtgttgaaacttgataacactatattccttga 667
Db 602 ACTTCGTGACCGAGCGCCGAGCGCCAGGAGCGGCTGCTGAGAGATCCTTACATTCGCTGG 661
Qy 668 atgaaaaaagattactagacatgaaagacatgtctacacatcttgaacaaatgtgtaaag 727
Db 662 TCAGCTCCAGAGTGTGACCGCTCAAGAGTGTGCTCCGCTGAGAGAGTCAATCCAGG 721
Qy 728 taaacgttcaactcctatattatgtctgaaagcgtagaaggtgaaagacttgcacacttg 787
Db 722 CCGGCAAGCCGCTGTGATTCCTGAGAGCGGAGCGAGCGGAGCGCCCTGTCACAGCTGG 781
Qy 788 tagtcaataaagctcgtgagagcaacttcaagtltgacgtaaaagcctctgtgttltgtg 847
Db 782 TGGTTCACAGATCCGCGCGCACCTTCAAGTCCGCTGCGGCTCAAGGCTCGGCGCTTCGGT 841
Qy 848 aacgcgtlaaagctatgtctgaaagataatgtctactcctactgaggaagaaacaaatattg 907
Db 842 ACCGCGGAGAGCATGTGACAGACATGCGCATTCCTCAGCGGTGATAGGTTCGTACAGG 901
Qy 908 aagatcgtgtatataaagcttgaagaaatgaaagctgtctctttaggaacagctaaagtg 967
Db 902 AAAGAGTCCGCTGTCCCTTGAGAGACCGCGGACGTCTCGCTGCGCCAGCGCCGCAAGG 961
Qy 968 tagtattgaaagaagaataactactatcgttgaatgtgtgtgtgaaataactaaagataata 1027
Db 962 TCGTCGTACCAAGAGACGAGACACATGCTGAGGCGCTCGGCGCATTCGATGCGCATCG 1021
Qy 1028 aagctcgtatlaaacaanaatctgtgcacaataatgaaagaacaaagctcagattatgtctg 1087
Db 1022 CCGGCGCGGTGTGATTCGCGCGGATTCGAGAACAGGAGCTCCGACATACGACCGCG 1081
Qy 1088 aaaaacttcaagaagcgttgcacaaactgtgtgtgaaagtagctgtatccatgttgag 1147
Db 1082 AGAAGCTTCAGACGCCCTTCGCCCAAGCTGCGCGCGGTGTTCGCTGATCAAGCGCGGAG 1141
Qy 1148 ctgtcaactgaaacttgaalagaagaagaagatcgtgtgaaagatgctcctaaatgtgaa 1207
Db 1142 CTGCCACCGAGGTGAGGCTCAAGAGAGCCGACACCGCATGAGAGCGCGCTCCGCAAGC 1201
Qy 1208 caaagctcgtgttgaagaagatattgtcctcgtgtgtgtgtatcttcttgcgtcca 1267
Db 1202 CGAAGGCTGCGCTGCAAGAGGCAATCCTGCGCGGTGCGGTGCTTGTGCGAGTGG 1261
Qy 1268 ttaaaagctcgtatgataataaactgtgtatgataatgaaacttgcgtgacttaataca 1327
Db 1262 CTCTGCGCTGAGACGACCTTCGCGCTGAGAGCGGCGAGGAGG---CCACGCGTGCACATCG 1318

Qy 1328 tccgtcgttctcttgaagaagccttactcaaatgtctgcaaatgtgtgtatgaaagtt 1387
Db 1319 TCCGGGTGCGCGCTGTGCGCGCTCCGCTCAAGCAAGATGCGCTTCAAGCGCGCGCTGGAGCGCG 1378
Qy 1388 ctattgtgtatgaaagaagttcgttgacacaaagaatgtgttggatcttaatgtctgacag 1447
Db 1379 GCGTGTGCGCGAGAGAGTGTTCACACCTGCGCGCGGTGACGCGGCTCAACCGCGCGAGCG 1438
Qy 1448 gagataagaagccttattaaagctgtgtgtatgtatgtatcctcaaaaagttaacgtattg 1507
Db 1439 GTGAGTACAGAGACCTGCTCAAGGCGCGGCTGCGCGCGCGTGAAGGTCAACCGCTGCG 1498
Qy 1508 catcaaaatgacagcatcagtagcctcctacttacttaactaagaagatgcgtatgtgtg 1567
Db 1499 CGCTGCAAGAGCGCGCGCTCATCGCGCTCTGTCTTCTTCAACCAAGAGCGCGCTGCGCG 1558
Qy 1568 aaaaaccaga 1577
Db 1559 ACAAGCCGGA 1568

RESULT 14

US-08-997-362-159
Sequence 159, Application US/08997362

Patent No. 5985287

GENERAL INFORMATION:

APPLICANT: Tan, Paul

APPLICANT: Hiyma, Jun

APPLICANT: Visser, Elizabeth

APPLICANT: Skinner, Margot

APPLICANT: Scott, Linda

APPLICANT: Prestidge, Ross

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR

TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS

NUMBER OF SEQUENCES: 194

CORRESPONDENCE ADDRESS:

ADDRESS: Law Offices of Ann M. Speckman

STREET: 2601 Elliott Avenue, Suite 4185

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98121

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/997,362

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/973,970

FILING DATE: June 12, 1997

APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347

FILING DATE: August 29, 1996

ATTORNEY/AGENT INFORMATION:

NAME: Sleach, Janet

REGISTRATION NUMBER: 37,007

REFERENCE/DOCKET NUMBER: 11000.1002c2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-269-0565

TELEFAX: 206-269-0563

TELEX:

INFORMATION FOR SEQ ID NO: 159:

SEQUENCE CHARACTERISTICS:

LENGTH: 1626 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-997-362-159

Query Match 23.7%; Score 390.8; DB 2; Length 1626;
 Best Local Similarity 53.98; Pred. No. 5e-94;
 Matches 847; Conservative 0; Mismatches 717; Indels 6; Gaps 2;

QY 8 cttaagaagaaccccttttgaatcgaagcccgtaaaacattcagagagtgatataac 67
 DB 5 CCAAGACAAATTGCCCTATGACGAGAGAGCCCGCCGCTCGAGCGGGCCCTCAACCCC 64
 QY 68 tgcgaatgctgttaaaagtaaacattggaacgaagccgtaaltgctatgtaaaagt 127
 DB 65 TCCGACAGCCGGTAAAGAGTACGTTGGCCCGAAGAGGTCCGACAGTGCCTGAGAGAGA 124
 QY 128 ctcttggtcccaagtattacaagaatggtgtatctgttgcgaagaatgtaactgt 187
 DB 125 AGTGGGCGCGCCGACCATACCAACGATGGTGTCTATCGCCCAAGAGATGACAGCTGG 184
 QY 188 aagaataagttgaataatgagcgctcaaatgtgttaagaagtgaagcccaaaactag 247
 DB 185 AGCAGCCCTACGAGAAATGCGCGCTGAGCTGCTCAAGACGTCGCCAAGACAGCAGC 244
 QY 248 alaligtgtgtatgtaaacatacaagcaacagtccttgcaagaactataltcgtgaag 307
 DB 245 ACGTCCGCGGCGAGCCGACCAACCGGCTCGCTGCTGAGGCTCTGTTGCGGAG 304
 QY 308 gtgtaaactgtgtagcagctgtgtcttaacctatgacctaaagtgatagataag 367
 DB 305 GCGTGGCCGAAAGTGCAGCGCGCGCCCAACCGCTCGGCTCAACGCTGAGATCAGAGAG 364
 QY 368 ctgtgtgtgtgttacttaagaagaactgaagcattacaagaactcgtgagcaaaag 427
 DB 365 CTGTCGAGGCTGTACCCACCATCGCTCTGAAAGTGGCCAAAGAGTGCAGACCAAGAGC 424
 QY 428 aaatagctcaagttggaacacattctgcaaaactgtatatacaaatagtaatactag 487
 DB 425 AGATTCTGCGACCGCGCGGATTTCCCGCGC--GACACCAAGATGCGGAGCTCAVCG 481
 QY 488 ctgaagctatggtcgaagttggaagaagtggtatcatcagttggaagaagctaaagtc 547
 DB 482 CCGAGGCGCATGACAAAGTGGGCAAGAGGTGTCAATCAGCTGAGAGTTCGAAACCT 541
 QY 548 ttgaactacattgaatgtgtgtaagaagttggaacgtggtgtaacctctccat 607
 DB 542 TCGGCGCTGCAAGCTGCACTGACGAGGATGCGCTTCGACAGAGGCTACATCTCGGCTT 601
 QY 608 acttltgaactaaacccctgaagaataatggttctggaactgtataacctatatactgt 667
 DB 602 ACTTCGTGACCGAGCGCGGAGCGGAGAGCGCTGAGAGATCCCTACATCTCTGCTGG 661
 QY 668 atgagaagaagatctacagatgaagaacatgttaccacactctagaacaagtgtctaag 727
 DB 662 TCACTCTCAAGGTGTCAAGCGTCAAGGATCTGCTCCGCTGCTGAGAAAGTCAATCCAGG 721
 QY 728 taacgcgtcaactccttattattgtctgaagaagctgaagaagctgtgcaacattg 787
 DB 722 CCGGCAAGCCGCTGCTATATATCCCGGAGAGCTGAGGGGAGGCGCCCTGTCACGCTGG 781
 QY 788 laagtaataagctcgtggaagcactccaagltgtagccgttaaaagctcctggttltgt 847
 DB 782 TGTGCAACAAGATCTCCGCGGACCTTCAAGTCCGCTCGCCGCAAGCTCCGCGCTTCGCTG 841
 QY 848 aagcggttaagctatgctgtgaagataatgttataccttactcgtgagaggaagcaatattg 907
 DB 842 ACCCGCCGAGGCGATCTGAGGACATGGCGATCTCACCGGTGTGAGTGTCTGTCAGCG 901
 QY 908 aagctgtgtatataagctgtgaagaatgtaagctgtcttcttggagacgttaaaagtg 967
 DB 902 AAAGAGTGGGCTCTCCCTGAGACGCCGCAAGCTCTGCTGCTGCGCCAGGCGCCGAGAG 961
 QY 968 laagtaataagcaaaalaactactatcgtltgaltgtgtgtggaagaatcgaagaata 1027
 DB 962 TCGTGTACCAAGAGAGAGACCAACATCTGTCGAGGCGCTCGGCGGATTCGATCCATCGC 1021

QY 1028 aagctgagttaaacaattcgtgcacaatttgaagaagaacagctcagattatgctg 1087
 DB 1022 CCGCGCGGTGCTCAATCGCGCGAGATCAAGAACAGCACTCGACTACGACCGC 1081
 QY 1088 aaaaacttcaagaagcttctgcaaaactgttgtgagtagctgtatccattggaag 1147
 DB 1082 AGAAGCTGCAAGAGCGGCTGCGCAAGCTGCGCGGCTGTTCGGTATCAAGCGCGAG 1141
 QY 1148 ctgctactgaactgaaatgaagaagaagaagatcgttgaagaatgctctaaatgcaa 1207
 DB 1142 CTGCCACGAGGTGAGACTCAAGAGAGCGCAAGCACCATCGAGAGACCGCTCGCAAGC 1201
 QY 1208 caagagctgtgttgaagaagatattgtccctgtgtgtgtgtgtgtgtgtgtgt 1267
 DB 1202 CGAAGCTGCGGTGGAAGAGGATCGCTCGCGGTGCGGTGCGGTGCTGCTGCTGCTG 1261
 QY 1268 ttaagctccttgaatataataaccgtgtatgatagtgaactgtcgtgactataatca 1327
 DB 1262 CTCTGCGGCTGAGAGAGACTCGGCTGACGCGGCGCAGAGG---CCACCGGTGCCAATCG 1318
 QY 1328 tccgtgtctcttgaagaagccttactacgttcaaatgtcgtgaatgtgtgtatgaag 1387
 DB 1319 TCCGCGTGGCGGTGCTGCGGCTCGCTCAAGCAGATCGCTTCAAGCGCGGCTGAGCGCG 1378
 QY 1388 ctatgtgtgaagaagaagctcgtgaaccaaagaatgttltgatttaagctgcatag 1447
 DB 1379 GCGTGTGCGGAGAGAGTGTCCAACTGCGCGCGGCTGAGGCTCAACGCGCGCGAG 1438
 QY 1448 gagaataagaaccccttataaagctgtgttcatatgtatccataaaagattacagta 1507
 DB 1439 GTGAGTACGAGAGACTGCTCAAGCGCGGCTCGCCGAGCGGAGTACCCGCTCGG 1498
 QY 1508 catcaaaaatgcaagcactagtagcctccttacttacttaactagaagaatgcgtatgt 1567
 DB 1499 CGGTGCAAGAGCGGCGCTCCATGCGGCTGCTGTTCTACACAGGAGCGGCTGCGCGC 1558
 QY 1568 aaaaaccaga 1577
 DB 1559 ACAAGCCGGA 1568

RESULT 15
 US-09-095-855-159
 ; Sequence 159, Application US/09095855
 ; Patent No. 6160093
 ; GENERAL INFORMATION:
 ; APPLICANT: Tan, Paul
 ; APPLICANT: Visser, Elizabeth
 ; APPLICANT: Skinner, Margot
 ; APPLICANT: Prestidge, Ross
 ; TITLE OF INVENTION: Compounds and Methods for
 ; NUMBER OF SEQUENCES: 208
 ; TITILE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Law Offices of Ann W. Speckman
 ; STREET: 2601 Elliott Avenue, Suite 4185
 ; CITY: Seattle
 ; STATE: WA
 ; COUNTRY: USA
 ; ZIP: 98121
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/095,855
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/705,347
 ; FILING DATE: 29-AUG-1996
 ; APPLICATION NUMBER: 08/873,970

FILING DATE: 12-JUN-1997
 APPLICATION NUMBER: 08/997,362
 FILING DATE: 23-DEC-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Sleath, Janet
 REGISTRATION NUMBER: 37,007
 REFERENCE/DOCKET NUMBER: 11000,1002c3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 206-269-0565
 TELEFAX: 206-269-0563
 TELETYPE: 206-269-0563
 INFORMATION FOR SEQ ID NO: 159:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1626 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-095-855-159

Query Match 23.7% Score 390.8; DB 4; Length 1626;
 Best Local Similarity 53.9% Pred. No. 5e-94;
 Matches 847; Conservative 0; Mismatches 717; Indels 6; Gaps 2;

Oy 8 ctataaagaatcccttttgaatgaagccgtgaaanaaacttcacagagtgatagaataac 67
 Db 5 CCAACACATTCGCTATGACGAGAGGCCGCCGCTGAGCGGGGCCCTCAACGCC 64
 Oy 68 ttgcgaatgctgttaagaataacacttgaaagccgtaatgcttatgaaagt 127
 Db 65 TCCCAGACCCCTTAAGGTGACGCTTGAAGGCCGCAAGGTCGCAACCTGCTGCTGAGAGA 124
 Oy 128 ctttggcttccagcttcttaacaaagatgctgttctgttgcgaagaatgtaactg 187
 Db 125 AGTGGGGGCCCCACGATTCACCAACGATGCTGCTCCTCCCAAGGATCGAGCTGG 184
 Oy 188 aagataaatttgaataatggtgcgtcaaatggttaagaagtgatcccaaaactacg 247
 Db 185 AGGACCCGTACGAGAGATCGGCGCTGAGCTGTCAAGAGGTGCCAAGAGACGAGCG 244
 Oy 248 atattgctgtgtgaagaaactaacacagcaacagtccttgacaaactatattatgaa 307
 Db 245 ACGTCGGGGGACGCCACACACCGCCACCGTGTGCTGAGCTGTGCTGCCGAG 304
 Oy 308 gttgaaactgtgagcaactgtgtaactatgacatgacatgaaagtgacatagaataag 367
 Db 305 GCTTCGCAACGTCGACCCGCGCAACCGCTGCGCTCAAGGCTGACATCGAGAGG 364
 Oy 368 ctgttgttctgttaactaaagaactaagacatatacaaaagcctactcgtgacaaag 427
 Db 365 CTGTGAGGCTGTACCCAGTGGCTGTGAGTGGGCCAAGAGGTGCGAGACCAAGAGCG 424
 Oy 428 aataagctcaagttgaaacattctgtcaaacctgtatatacaaaatgaatgataag 487
 Db 425 AGATTTCGACCCCGCGCGATTTCCCGCGC---GACACCCAGATCGCGAGCTCATCG 481
 Oy 488 ctgaagctatagctaaagtltgaaagaaggtgtatatacagttgaaagaagtaagtc 547
 Db 482 CCGAGGCCATGAGCAAGGTCGCAAGAGGCTGATCATCGTCAAGGATGGAACACCT 541
 Oy 548 ttgaactaatatagatgtgtgtgaagaagtaagttgaaacgtgtgtaactcctcat 607
 Db 542 TCGGCGCTGAGCTGAGCTCAACGAGGATGCTGCTGAGCAAGGCTACATCTGGGTT 601
 Oy 608 acttgaactaaatccctgaagaatgtgtgtgaactgtataacctatatacttctgta 667
 Db 602 ACTTCGTACCGACCCCGACGCCAGAGGCCGCTCTGAGAGATCCCTACATCTGCTGG 661
 Oy 668 atgaagaataaatactagcaatgaagaatgatacaaatcttagaacaagttgtaag 727
 Db 662 TCAGCTCCAGAGTGTGACCGTCAAGAGATCTGCTCCGCTGAGAGAGGTCAATCGAG 721
 Oy 728 taacacgtcaactcttatattgtgaagacgtagaaggtgaagcaacttgtaactg 787

Db 722 CCGGCAACCCGCTGCTGATCATCGCCGAGGAGCTGAGCGGAGGCCCTGTCCACGCTGG 781
 Oy 788 tagtcaataaagctccgttgagagacacccaagttgtatgacggtlaaagctcgtgttgg 847
 Db 782 TGTCTACAGATCCGCGCACCTTCAAGTCCGTCGCTCAAGGCTCCGGGCTTCGGTG 841
 Oy 848 aacgcgtlaaagctatgcttgagaatattgtatctacttactgtgaagaagaataattg 907
 Db 842 ACCGCGCAAGCGCATGCTGCGAGCATGCGCATCTCAACCGTGTGTGAGTGTGTCAGCG 901
 Oy 908 aagatcgtgtataaagcttgaaatgtgaagctgtctctttagaaacagtaaacgtg 967
 Db 902 AAAGAGTGGCGCTTCCTCGAGACCGCGAGCTGTCTCTCTGCGCAGGCCCGCAAGG 961
 Oy 968 tagtattgacaaagaataactactatcgttgaatggtgtgtgtgaataatcagaatatta 1027
 Db 962 TGTCTGTCACCAAGAGCAGACACCATCTGTCGAGGCTTCGGGCGATTCGATCCATCG 1021
 Oy 1028 aagctgagttaaanaattcgtgcacaatttgaagaagaacagctcagattatgactg 1087
 Db 1022 CCGGCGGCTGCTCAGATCCCGCGAGATGAGAACAGCATCCGACTACGACCGCG 1081
 Oy 1088 aaaaacttcaagaacgtcttgaaactgtgtgtgtgaagtgttataccaagtgtgag 1147
 Db 1082 AGAAGCTCAGAGAGCGCTGCGCAAGCTGCGCGGCTGTGCTGATCAAGGCCGAG 1141
 Oy 1148 ctgctactgaaacttgaatgaagaagaagatcgtgtgtaagaatgctcttaattgcaa 1207
 Db 1142 CTGCCACCGAGGTGAGCTCAAGAGAGCGCAACCGCATCGAGACCCGCTGCCAAG 1201
 Oy 1208 caaagctgctgtgaagaagttatgtccctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1267
 Db 1202 CGAAGGCTGCTCCAAAGAGGAGCATGCTGCGGAGGCGGCTGCTGCTGCTGAGTGG 1261
 Oy 1268 ttaaaagcttgaatgatatlaaacctgtgtatgaatgaatgactgtgacttaataca 1327
 Db 1262 CTCTGCGCTGTGAGACCTCGGCTGACGCGGCGGAGAGG---CCACCGGTGCCAATCG 1318
 Oy 1328 tccgtcgttctcttgaagaagccttgaatgaatgtcgtgaatgtcgtgtgtgtgtgt 1387
 Db 1319 TCCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1378
 Oy 1388 ctatgtgtgtgaagaagcttgaacaaagaagatgtgtgtgtgtgtgtgtgtgtgtgtgt 1447
 Db 1379 GCGTGTGTGCGAGAGGTGTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1438
 Oy 1448 gagaatatgaagaaccttataaagctgtgtgtatgtatccttaaaaaagttacagttatg 1507
 Db 1439 GTGACTACGAGACCTGCTCAAGCGCGGCTGCGGACCGGTGAAGGTCAACCGCTGG 1498
 Oy 1508 cattacaataatgacatcagtaagctccttaacttactaactaagaatgactatgtgt 1567
 Db 1499 CGCTCAGAACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1558
 Oy 1568 aaaaacaga 1577
 Db 1559 ACAAGCCGA 1568

Search completed: August 1, 2001, 15:59:11
 Job time: 1772 sec

Thu Aug 2 08:04:53 2001

us-09-077-574a-1.rni

Page 18

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2001, 15:28:08 ; Search time 1278.56 Seconds

(without alignments)
12176.863 Million cell updates/sec

Title: US-09-077-574A-1
Perfect score: 1647

Sequence: 1 atgctctcaagaatacct.....gtatggacgtagtactag 1647

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 segs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est33:*
111: gb_est34:*
112: gb_est35:*
113: gb_est36:*
114: gb_est37:*
115: gb_est38:*
116: gb_est39:*

117: gp_est48:*
118: gp_est49:*
119: gp_est50:*
120: gp_est51:*
121: gp_est52:*
122: gp_est53:*
123: gp_est54:*
124: gp_est55:*
125: gp_est56:*
126: gp_est57:*
127: gp_est58:*
128: gp_est59:*
129: gp_est60:*
130: gp_est61:*
131: gp_est62:*
132: gp_est63:*
133: gp_est64:*
134: gp_est65:*
135: gp_est66:*
136: gp_est67:*
137: gp_est68:*
138: gp_est69:*
139: gp_est70:*
140: gp_est71:*
141: gp_est72:*
142: gp_est73:*
143: gp_est74:*
144: gp_est75:*
145: gp_est76:*
146: gp_est77:*
147: gp_est78:*
148: gp_est79:*
149: gp_est80:*
150: gp_est81:*
151: gp_est82:*
152: gp_est83:*
153: gp_est84:*
154: gp_est85:*
155: gp_est86:*
156: gp_est87:*
157: gp_est88:*
158: gp_est89:*
159: gp_est90:*
160: gp_est91:*
161: gp_est92:*
162: gp_est93:*
163: gp_est94:*
164: gp_est95:*
165: gp_est96:*
166: gp_est97:*
167: gp_est98:*
168: gp_est99:*
169: gp_est100:*
170: gp_est101:*
171: gp_est102:*
172: gp_est103:*
173: gp_est104:*
174: gp_est105:*
175: gp_est106:*
176: gp_est107:*
177: gp_est108:*
178: gp_est109:*

190: gp_est110:*
191: gp_est111:*
192: gp_est112:*
193: gp_est113:*
194: gp_est114:*
195: gp_est115:*
196: gp_est116:*
197: gp_est117:*
198: gp_est118:*
199: gp_est119:*
200: gp_est120:*
201: gp_est121:*
202: gp_est122:*
203: gp_est123:*
204: gp_est124:*
205: gp_est125:*
206: gp_est126:*
207: gp_est127:*
208: gp_est128:*
209: gp_est129:*
210: gp_est130:*
211: gp_est131:*
212: gp_est132:*
213: gp_est133:*
214: gp_est134:*
215: gp_est135:*
216: gp_est136:*
217: gp_est137:*
218: gp_est138:*
219: gp_est139:*
220: gp_est140:*
221: gp_est141:*
222: gp_est142:*
223: gp_est143:*
224: gp_est144:*
225: gp_est145:*
226: gp_est146:*
227: gp_est147:*
228: gp_est148:*
229: gp_est149:*
230: gp_est150:*
231: gp_est151:*
232: gp_est152:*
233: gp_est153:*
234: gp_est154:*
235: gp_est155:*
236: gp_est156:*
237: gp_est157:*
238: gp_est158:*
239: gp_est159:*
240: gp_est160:*
241: gp_est161:*
242: gp_est162:*
243: gp_est163:*
244: gp_est164:*
245: gp_est165:*
246: gp_est166:*
247: gp_est167:*
248: gp_est168:*
249: gp_est169:*
250: gp_est170:*
251: gp_est171:*
252: gp_est172:*
253: gp_est173:*
254: gp_est174:*
255: gp_est175:*
256: gp_est176:*
257: gp_est177:*
258: gp_est178:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| DB | 715 | ATACCTTTATTAAATACATCAAAAGGTCAGAAATCTGAAATTCACAGATGCCATATCTTCGT | 774 |
|-----------------------|--|---|----------------------|
| Oy | 666 | taatgagaaaaagatctacgacatgtaagaacatgctaccatcttgaaacaagtgcgtaa | 725 |
| Db | 775 | GAGTGAAGAAAGAAATTTCTAGTATCCAGTCCCAATTGACTGCTCTTGAAATTCACAAATGC | 834 |
| Oy | 726 | ggttaaacgcgtcacccctcttatctgctgtgaagaacgtgaagagagagagcacttgcacact | 785 |
| Db | 835 | TCACCGTAGAGCCCTTTGGTATATATCCCTCAAGATGTTGATGAGGAAGCTTAAGTACACT | 894 |
| Oy | 786 | tgtatgcaataaagctccgtggaagcactcccaagcttgtagcgttaaaagctcctggtcttg | 845 |
| Db | 895 | CCTCTTGAAATAGCGCTAAAGGTTGCTCTTACAGTTGTCGCGCAGTCAAGCCCTCCAGGCGTTTG | 954 |
| Oy | 846 | lgaacgcgcgtlaaagctatgcttgaagatalgtctatccctactgcgtgaagaga | 896 |
| Db | 955 | GTGACAAATAGAGAGAGAGAGAGATGAGATGAGTGTGATGAGGAGAGCTTAAGTACACT | 1005 |
| RESULT | 2 | | |
| LOCUS | BF275584 | 878 bp | mRNA EST 07-MAR-2001 |
| DEFINITION | GA__EB0024G23f Gossypium arboreum 7-10 dpa fiber library Gossypium | | |
| ESION | arboreum cDNA clone GA__EB0024G23f, mRNA sequence. | | |
| KEYWORDS | BF275584 | | |
| SYNON | BF275584.1 | GI:11206654 | |
| ORGANISM | EST. | | |
| SOURCE | Gossypium arboreum. | | |
| REFERENCE | Gossypium arboreum. | | |
| AUTHORS | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | |
| TITLE | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; | | |
| JOURNAL | Rosidae; eurosids II; Malvales; Malvaceae; Gossypium. | | |
| COMMENT | 1 (bases 1 to 878) | | |
| | Wing,R.A., Flisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry | | |
| | ,D., Wood,T.C., Leslie,A. and Wilkins,T.A. | | |
| | An integrated analysis of the genetics, development, and evolution | | |
| | of the cotton fiber | | |
| | Unpublished (2000) | | |
| | Contact: Wing RA | | |
| | Clemson University Genomics Institute | | |
| | Clemson University | | |
| | 100 Jordan Hall, Clemson, SC 29634, USA | | |
| | Tel: 864 656 7288 | | |
| | Fax: 864 656 4293 | | |
| | Email: rwing@clemson.edu | | |
| | Seq primer: TAATACGACGACGATGATGAG | | |
| | High quality sequence stop: 716. | | |
| FEATURES | location/qualifiers | | |
| source | 1..878 | | |
| | /organism="Gossypium arboreum" | | |
| | /strain="AKA" | | |
| | /cultivar="8400" | | |
| | /db_xref="taxon:29729" | | |
| | /clone="GA__EB0024G23f" | | |
| | /clone_lib="Gossypium arboreum 7-10 dpa fiber library" | | |
| | /tissue_type="Fibers isolated from bolls harvested 7-10 | | |
| | dpa" | | |
| | /lab_host="E. coli" | | |
| | /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI" | | |
| BASE COUNT | 253 a 157 c 219 g 245 t 4 others | | |
| ORIGIN | | | |
| Query Match | 19.0%; Score 312.2; DB 146; Length 878; | | |
| Best Local Similarity | 61.8%; Pred. No. 7.9e-69; | | |
| Matches | 494; Conservative 0; Mismatches 306; Indels 0; Gaps 0; | | |
| Oy | 462 | tgatacaacaataggtaatacatcatagctgaagctatgcttaaaagtctggaagagagtgct | 521 |
| Db | 17 | TGATGACCTTATGAGAACATGATGATGCTGATGCAATTCACAAAGTTGACCTGATGCTGT | 76 |
| Oy | 522 | tatcaacgttgaagaaagctaaagctcttgaaactcatctagatgtgtgttgaagaaatgaa | 581 |

| | | | | |
|------------|---|--|-------------------------------------|--------------------------|
| Db | 77 | TTTGTTCATGATGATCATCATCTCTCTCA | TTTGTGAGCACCACTTGATGTTGAGGAAGAATGCA | 136 |
| Qy | 582 | gtttgacccgtggacatcctctccatctctactctgttaactaatccctgagaagaatggtttgga | 641 | |
| Db | 137 | GATTGTACAGAGGTTTACATCTCCCAACAAATTTTGTATTAACCCGGGAATATGTTTGGA | 196 | |
| Qy | 642 | actttgataaccttatatcctcttctgtgataatgagaaaagatctactgacatgaaagacatgct | 701 | |
| Db | 197 | GTTTGTGATGATGCAAGAGTGTGTTGGTAACAGATGATCAAAAGATTTCAGCTATTAAGACATCAT | 256 | |
| Qy | 702 | accacatcttagaacacaagttgtctaaagtgtaaacccgtccactccttatatctgtgaagacgt | 761 | |
| Db | 257 | TCCCGCTGTTAGAAAGAACCCACCAATTAAGATTCCTCTCTTGTATTAATTTGCTGAGGATGT | 316 | |
| Qy | 762 | agaaggtgaaagactctgcaacactgtgtatgtaataagctccgttggagaagaccgaattgtc | 821 | |
| Db | 317 | GAGTGGAGAGGCTCTGGCCACACCTTGTGTGTACAACTGGCTGGCATTTCCAATGTGTGC | 376 | |
| Qy | 822 | agccgtlaaaagctctctggttttgtgtgaaagccgcgttaagactatgtcttgaagaatctgtcat | 881 | |
| Db | 377 | ACCCTATTAAAGCTCTGGTGTGTTGGTAAGGAAGGAAGAAAGCTCTCCTTCANAGATATTGCCAT | 436 | |
| Qy | 882 | cccttcctgagagaagaagacaatatcttgaagatcgtgtgtataagcttgaagcttgaagaatg | 941 | |
| Db | 437 | TCTGACTGCTGTGAGTTCCAAAGCTAGTGAATTTGGTGTCTGCTGCGAATPACCCTCAGT | 496 | |
| Qy | 942 | gtctctctttagaagacagctttaaagcgtgtatgttatctgacaaagaaaataactactatcgttga | 1001 | |
| Db | 497 | TGAGACAGCTTGTGATATTGCCAGAAAGGTATATCTTACCAAGATTCGACTCAACTAATTTGC | 556 | |
| Qy | 1002 | tgggtctggaagaatcagaagatatataagctcgagatgaataaactcgtgcacaaacttga | 1061 | |
| Db | 557 | TGAMCACCCTCAAAAGATGAGATACANAAGTAGGCTGACACAACTAATAAAAAGAGCTAGC | 616 | |
| Qy | 1062 | agaacaagctcagatctatgatctgtgaaaaaactccaagaagacgtcttgcaaaacttgttg | 1121 | |
| Db | 617 | TCAGACACATTTCTGCTATTGTGATTCAGAAAAATTTGGCACAAAGATTTGCCAAATCATCTGG | 676 | |
| Qy | 1122 | tggaaatagctcttaccatctgtctgagcgtctgacacgaagaacatgaagaagaagaagga | 1181 | |
| Db | 677 | TGGTTTTCAGATCATTAAGGTGGGGGCTGCAACAGAGACTAACTTGNGATTCGTAAAGCT | 736 | |
| Qy | 1182 | tcgtgtagaagaatgctctaaatgacaacaagagctcggttgaagaagaagtatgtccctgg | 1241 | |
| Db | 737 | ACGGATGGAAGATGCAAGAAGATGCTACACTTTGCTGCCATGMAAAGATATTGCTCGG | 796 | |
| Qy | 1242 | tgggtgtactcgtcttctgtcc | 1261 | |
| Db | 797 | TGGTGTGCTGCTGCTTGAATTC | 816 | |
| RESULT 3 | | | | |
| AL515579 | LOCUS | AL515579 | 1017 bp | mRNA |
| DEFINITION | AL515579 | LTI_NF1011.NBC1 | Homo sapiens | CDNA clone CS0D002YD20 5 |
| ACCESSION | AL515579 | | | prime_rRNA sequence. |
| VERSION | AL515579.1 | GI:12779072 | | |
| KEYWORDS | EST. | | | |
| SOURCE | human. | | | |
| ORGANISM | Homo sapiens | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | |
| AUTHORS | Li, W.B., Gruber, C., Jasee, J. and Polayes, D. | | | |
| TITLE | Full-length cDNA libraries and normalization | | | |
| JOURNAL | Unpublished (2001) | | | |
| COMMENT | Contact: Genoscope | | | |
| | Genoscope - Centre National de Sequencage | | | |
| | BP 191 91006 EVRY cedex - France | | | |
| | Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr. | | | |
| FEATURES | Location/Qualifiers | | | |
| source | 1..1017 | | | |


```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DA002YD20"
/clone_1ib="J71_NFL001_NBC1"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: PCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the PCMVSPORT 6
vector. Library is not normalized, but is the control for
the normalized libraries. Library was constructed by Life
Technologies. Contact: Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
, Maryland 20850, USA Fax: (1) 301 610 8371 Email:
f.liang@lifestech.com URL:
http://fulllength.invitrogen.com"
BASE COUNT      308 a      191 c      255 g      261 t      2 others
ORIGIN

```

```

Query Match      18.8%; Score 310.4; DB 105; Length 1017;
Best Local Similarity 60.0%; Pred. No. 2.4e-68;
Matches 534; Conservative 1; Mismatches 352; Indels 3; Gaps 1;

```

```

QY      6  ttctaaagaatcccttttgatgctaaagcccggtgaataacttcacgaggtgtagataa 65
DB      115  TCCCAAGATGTAATAATTTGGTGCACATGCCAGCCCTTAATGCTTCAAGGTGACACCT 174
QY      66  actgcaaatgctgttaagaacttgaccctaaagccgtaatgctgctatgataa 125
DB      175  TTACGCCGATGCTGTGGCGCTTACATGGCCCAAGACAGACAGTATTTAGCA 234
QY      126  gctcttggtccccaagtattacaagaatggtgctatcgttgcaagaatgtaact 185
DB      235  GAGTTGGGGAAGTCCCAAGTAACAAGATGCTGACTGTGCAAGTCAATTGACTT 294
QY      186  tgaagaataatgtaagaataatggtgctaaatggttaagaatgctgctcccaactg 245
DB      295  AAAAGTAAATACAAAAAGTTCGAACTTAACCTTGAAGATGTTGCCAATGACACAA 354
QY      246  cgaatctgctggtgtagaactacacaaacgaactccttgcaagaactatcctgta 305
DB      355  TGAAGAGCTGGGATGCGACATCCACTCTACTGACTGCGACAGCTCTATAGCCAGCA 414
QY      306  aggtgtaaaactgttagcagctggtcgttaactcctatgccaatgacgtgcatagaa 365
DB      415  AGGCTTCGAGAGATTAGCAAGGTGCTAATCCAGTGAATACAGAGAGGTGTATTT 474
QY      366  agctgtgtgtgctgttactaaagaactaagcgacatlaacaagctcactcgtgaccaa 425
DB      475  ACCCTGTTGATGCTGATTTGCTGAACCTTAAAGCAGTCTTAACCTGTGACCAACCCCT 534
QY      426  agaaatagctcaagttggaacatctctgcaactctgtatcacacatagatgtaata 485
DB      535  AGAAATGACAGAGCTTGCATACATTTCTGCAACGACAGCAAGAAATTTGCCATATTCAT 594
QY      486  agctgaagctatggtcgaagaatggtgctatcacagttggaagctgaagaag 545
DB      595  CTTCTGATGCAATGAAAGAGTTGGAAGAGGTGTCATCACAGTAAAGATGGAAGAAAC 654
QY      546  tctggaactacatagatggtgctggaagaatggaagttgacggtgctacctctcc 605
DB      655  ACTGATGATGATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 714
QY      606  atacttgaactaactccgagaagaatggttctggaacttgtaacctatatactcttg 665
DB      715  ATACTTTATTTATACATCAAGAGTCAGAAATGTCAGATTCAGATGCTTATGCTGTT 774
QY      666  taatggaagaagaatcactacgaagaagacatgctacacatcttagaagaatggtctaa 725
DB      775  GAGTGAAGAAAGAAATTTCTAGTATCCAGTCAATTGACTGCTGTGAATATTGCCAATGC 834

```

```

QY      726  agtaaacgcctccactcttatattgtctgaagcgttagaaggtgaagccacttgcacact 785
DB      835  TCACCCCTAAGCCCTTTGGTATATTCCTGAAGATGTTGAGGAGAGCTCTAAGTACACT 894
QY      786  tgaagcaataagctcgtggaagcactccaagttgtagccgtaagaagccctggtttg 845
DB      895  CGCTTGATTAAGCTTAAGAGTGTGCTTCAAGTGTGGCAGTAAAGCTCCAGAGGTTTG 954
QY      846  tgaagcgcgaagaagctatgcttgaagatatgtctacactactggaagag 895
DB      955  TGA---CATGAAGAAGACAGCTTAAGATATGCTTATGCTACTGCTGCTGCTG 1001

```

```

RESULT      4
AL532233      988 bp      mRNA      EST      13-FEB-2001
LOCUS      AL532233 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DM004YH20 5
DEFINITION      prime, mRNA sequence.
ACCESSION      AL532233
VERSION      AL532233.1 GI:12795726
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 988)
AUTHORS      Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
JOURNAL      Full-length cDNA libraries and normalization
COMMENT      Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

```

```

FEATURES
source
location/Qualifiers
1..988
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DM004YH20"
/clone_1ib="J71_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: PCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the PCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact: Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax: (1) 301 610
8371 Email: f.liang@lifestech.com URL:
http://fulllength.invitrogen.com"

```

```

BASE COUNT      306 a      188 c      242 g      251 t      1 others
ORIGIN

```

```

Query Match      18.7%; Score 308.8; DB 106; Length 988;
Best Local Similarity 60.3%; Pred. No. 6e-68;
Matches 528; Conservative 0; Mismatches 347; Indels 1; Gaps 1;

```

```

QY      6  ttctaaagaatcccttttgatgctaaagcccggtgaataacttcacgaggtgtagataa 65
DB      114  TCCCAAGATGTAATAATTTGGTGCAGATGCCAGCCCTTAATGCTTCAAGGTGACACT 173
QY      66  actgcaaatgctgttaagaacttgaccctaaagccgtaatgctgctatgataa 125
DB      174  TTACGCCGATGCTGTGGCGCTTACATGGCCCAAGACAGATGATTTATTTAGCA 233
QY      126  gctcttggtccccaagtattacaagaatggtgctatcgttgcaagaatgtaact 185
DB      234  GAGTTGGGGAAGTCCCAAGTAACAAGATGCTGACTGTGCAAGTCAATTGACTT 293

```

```

OY 186 tgaagataagcttggaaatgagcgcctcaaatggttaagaagtagctcccaaaactag 245
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 294 AAAAGATAAATACAAAACATTGGAGCTAAACTGTTACAGATGTGCAATTAACACAAA 353
OY 246 cgaatagctgtgtagtgaactacaagcaagcagctctgcacaaactatattcgtga 305
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 354 TGAAGAACCTGGGAGTGGCAGCTACACAGCTACTGCTACTGCTACTATAGCCAAAGCA 413
OY 306 agggtaaaacttgtagaagctgtgtagtgaactgagcctatgagccttaagctgagataa 365
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 414 AGGCTTCGAGAGATTACCAAGAGTCTTAATCCAGTGAATAACGAGAGGTGTGATGTT 473
OY 366 agctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 425
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 474 AGCTGTGATGCTGTAAATGCTGAACCTTAAGACAGCTTAACCTGTGACACCCCTGA 533
OY 426 agaaatagctaaagtgtgaaccattctgcacactctgatacaacaatagtaatacat 485
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 534 AGAATATGCACAGCTTCTACGATTTCTGCAACGAGACAAAGAAATTCGCAATATCAT 593
OY 486 agctgaagctatggtctaaagtgtgaagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 545
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 594 CTCTGATGCAATGAAAAAGTTGCAAGAGAGGTGTGCTACAGTAAAGGATGGAAGAAC 653
OY 546 tcttgaactacatagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 605
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 654 ACTGAATGATGATTAAGAAATTATTGAAGCATGAGATTGATGAGAGCTTATTTCTCC 713
OY 606 atacttgttaactaaccctgagagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 665
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 714 ATACTTTATTAATATCATCAAAAGCTCAGAAATGTGAATTCAGATGCTCTATGTCTGT 773
OY 666 taatgagaaaaagatlaactagcatgaaagacatgctacccaactcttaagaacagtgtctaa 725
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 774 GAGTGAAGAAAGAAATTTCTAGTATCCAGTCCATTGTAACCTGCTTTAAATTTGCCAATGC 833
OY 726 agtaaacctgcacactctcttatttctgtgaagcgttagaagcgttgaagcacttgaacact 785
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 834 TCACCGTAAGCCTTTGGTCAATAATCCGCTGAAGATGTGATGAGAGAACCTTAAGTACACT 893
OY 786 tgtatgtaataaagctccgtgtgagacatcccaagttgttagcgcgttaaaagcctctgttttg 845
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 894 CGTCTTCAATAGCTTAAGGCTTGTCTTCAGGTTGTGCGAGTCAAGCTCCAGCGTTTGC 953
OY 846 tgaagcgcgtlaaagctatgcttgaagaatgtgcat 881
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 954 TG-ACAAATAGAAAGAACGCTTAAGATATGCTAT 988

```

```

RESULT 5
AL515262 955 bp mRNA EST 13-FEB-2001
LOCUS AL515262 L11_NFL006_PL2 Homo sapiens cDNA clone CLOBB0162A04 5
DEFINITION AL515262 L11_NFL006_PL2 Homo sapiens cDNA clone CLOBB0162A04 5
VERSION AL515262
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 955)
AUTHORS Li, W.B., Gruber, C., Jessup, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

```

```

FEATURES
Source 1..955
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

```

/cclone="CLOBB0162A04"
/cclone_id="L11_NFL006_PL2"
/issue_type="placenta"
/note="Vector: PCMVSPORT 6; Site: 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the PCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@life.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 288 a 161 c 236 g 247 t 3 others
ORIGIN

```

```

Query Match 18.64; Score 306.6; DB 105; Length 955;
Best Local Similarity 60.14; Pred. No. 2.2e-67;
Matches 507; Conservative 1; Mismatches 335; Indels 0; Gaps 0;

```

```

OY 6 tcttaagaataatccctttagtgaagagccgtgaaataacttcaagagtgtagataa 65
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 113 TCCCAAGAGTGAATAAATTTGGTGCAGATGCCCGACCTTAATGCTTCAAGGTATGACT 172
OY 66 acttgcataatgctgttaagtaacacttgaagcgttaatgtcgttattgtaaa 125
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 173 TTAGCCGATGCTGTGCGCTTCAATGAGGCGCCAAAGGAAACAGTGAATTTATGAGCA 232
OY 126 gtcttgtgtcccgatlaacaaagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 165
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 233 GAGTGGGAGAGTCCCAAGTAAACAGTAAAGTGTGTACTATTTGCAAGTCAATGACTT 292
OY 186 tgaagataagtttgaataatgagcgtcaaatgttgaagaagtagtcccaaaactag 245
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 293 AAAAGATAAATACAAAACATTGGAGCTAAACTGTTCAAGATGTGCTCCATTAACACAA 352
OY 246 cgaatagctgtgtagtgaactacaagcaagcagctctgcacaaactatattcgtga 305
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 354 TGAAGAACCTGGGAGTGGCAGCTACACAGCTACTGCTACTGCTACTGCTATATGCAAGGA 412
OY 306 aggtgtaaactctgtgagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 365
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 413 AGGCTTCGAGAGATTAAGCAAGAGTGTCAATCCAGTGAAGATGAGAGCTGTGATGTT 472
OY 366 agctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 425
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 473 AGCTGTGATGCTGTATTGCTCAACTTAAGAGCACTTAACCTGTATACACCCCTGA 532
OY 426 agaaatagctcaagttggaaccattctgcaaacctgtatatacaaatagtaatacat 485
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 533 AGAATATGCACAGCTTGTCAAGATTTCTGCAACGAGAACAAAGAAATTTGCAATATCAT 592
OY 486 agctgaagcgtatgagcgttgaagaagaggtgtatataagcttgaagaagcgtatgagcgt 545
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 593 CTCTGATGCAATGAAAAAATTGCAAGAGAGGTGTCTATCAGCTTAAGATGGAAGAAC 652
OY 546 tcttgaactacatagatgtgtgtgaagaatgttgaagcgtgtgactctctc 605
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 653 ACTGAATGATGATTAAGAAATTATTGAAGCATGAGATTGATGAGAGCTATTTCTCC 712
OY 606 atacttgttaactaaccctgagagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 665
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 713 ATACTTTATTAATATCATCAAAAGCTCAGAAATGTGAATTCAGAGATGCTATGTTCTGTT 772
OY 666 taatgagaaaaagatlaactagcatgaaagacatgctacccaactcttgaagaagtgtctaa 725
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 773 GAGTGAAGAAAGAAATTTCTAGTATCCAGTCCATTGATGCTTGTGAGAGTCCCAATGC 832
OY 726 agtaaacgctcaactcttatttctgtgaagcgttagaagtgtaagcacttgaacact 785
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 833 TCACCGTAAGCCTTTGGTCAATATCCGTAAGATGTTGATGGAAGAGCTTAAGTACACT 892

```



```

Db      361  GTGAGAAATGTGAAATCCAGAGATGCTATGTTCTGTGAGGAGAAAGAAATTTCTAGTA 420
Qy      689  tgaagacatgctaccacatcttaagacaagttgctaaagttaaccgtccactcttacta 748
Db      421  TCCAGTGCATGTGACCTGCTCTTGAATTCGCAATGCTCACCAGTAAAGCTTTGGTCATAA 480
Qy      749  ttgttgagacgttagaaggtgaagcaacttgcaacacttgtaagtcacgtcgtagag 808
Db      481  TCCGTGAAGATGTTGATGAGACAGAACCTTAAGTACACTGCTTGTGAATAGCTTAAGGTTG 540
Qy      809  caaccacagttgtagccgtaaaagctccggttcttgtagaacgcccgaagctatgcttg 868
Db      541  GTCTTCAAGGTGTGCGACAGTCAAGGCTCCAGGTTGGTGAATAGAAAGAACAGCTTA 600
Qy      869  aagatctgctacatccctactctgtagagagaagaacatattggaagatcgtgta---taagc 925
Db      601  AAGATATGGCTATGTGCTACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy      926  ttgaagaatgtagcctgctctctttaggaacagctaaagctgtagtaattgacaagaag 985
Db      661  TTGAAGAGCTTCACGCTTCATGCTTACCTTAGGAAAGTTGGAGAGCTCATTTGTACCAAGACG 720
Qy      986  ataactactcgtttagttagtctggaataatcagaagaatataaagctcgaagctaaacaa 1045
Db      721  ATGCGATGCTCTTAAGAGAAAGAGTGAAGGCTCAAAATTTGAAAGAAAGCTATTCAGAA 780
Qy      1046  ttggtgcaacaatggaagaacagctcagattatgacgtggaagaacttcaagaagctc 1105
Db      781  TCATTGACGAGTATGATGCTCAGCACTAGTGAATGAAAGAAAGAAAGTGAATGAAAGGCG 840
Qy      1106  ttgcaaaactctgttgtagtagctgtatccatctgtagagctgctgaagctgaagctgaa 1165
Db      841  TTGCAAAAGCTTTCAGATGAGAGTGGCTGTCTGTAAGGTTGGTGGCAAGTATGTTGAAG 900
Qy      1166  tgaagaagaagaaga 1181
Db      901  TGAATGAAGAAARAAGA 916

RESULT  9
Bg599254      790 bp      mRNA      EST      12-APR-2001
LOCUS      Bg599254      csts Solanum tuberosum cdna clone csts2f20 5' sequence,
DEFINITION      mRNA sequence.
ACCESSION      Bg599254
VERSION      Bg599254.1 GI:13619095
KEYWORDS      EST.
SOURCE      Solanum tuberosum
ORGANISM      Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Eudicotyledons; Core eudicots;
Asteridae; easterids I; Solanales; Solanaceae; Solanum.
REFERENCE      1 (bases 1 to 790)
AUTHORS      van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Chleminco, A.,
      Bougri, O., Buell, C. R., Roeding, C., Tanksley, S. and Baker, B.
      Generations of ESTs from sprouting potato eyes
      Unpublished (2000)
JOURNAL
COMMENT      The Institute for Genomic Research
      For clone info: please contact Research Genetics, Libraries
      Division tel 1-800-711-6195, email cdna@resgen.com
      Seq primer: M13F-R.
FEATURES
SOURCE
      location/Qualifiers
      1..790
      /organism="Solanum tuberosum"
      /cultivar="Kennebec"
      /db_xref="taxon:4113"
      /clone="csts2f20"
      /clone11b="csts"
      /tissue_type="sprouting eyes from tubers"
      /dev_stage="12-14 weeks post harvest"
      /lab_host="SOLR"

```

```

BASE COUNT      254 a      130 c      212 g      194 t
ORIGIN
Query Match      17.4%; Score 286.8; DB 155; Length 790;
Best Local Similarity 60.3%; Pred. No. 2.3e-62;
Matches 474; Conservative 0; Mismatches 312; Indels 0; Gaps 0;

/note="Vector: pBluescript SK(-); Site-1: EcoRI; Site-2:
XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."
Qy      342  ggcaataacgttgacatagaataagctgttgctgttcaagaagaacgacat 401
Db      2      GGACCTTAGAGCGGGGTATTACCATGGCTGTAGATTCTGTGGCAAAACCTGAAAGCAG 61
Qy      402  tacaagcctactcgtgacccaagaataagctcaagttggaacattctgcaactc 461
Db      62  AGCAGGATGATCACTACATCATGAGCAAAATTCACAGCTGGGACTATCTCTGCAAAACG 121
Qy      462  tgataacaataagtaataatcatatgctggaagctatgctgaagttagaagaagagtg 521
Db      122  AGAAAGAGAAATTTGGCGACCTAATTGCAAGGCTTGAAGAGAAATGACCAAGCAGGAGT 181
Qy      522  taticagctgaggaagcctgaagctctgaaactacattagaatgctgtagaagaatga 581
Db      182  CATCACTATTCAAGATGGGAAACCTTCTCATATGAGTTGATGATGTTGTAAGGAGATGA 241
Qy      582  gttgaccgtgctacccctctccactacttgtaactaactcctggaagaatggttgtag 641
Db      242  GTTGATGATGAGGTTACATATCCATACCTCATCTCAATATGAGAAATCAGAAATCA 301
Qy      642  actgataaccttatccctcttgtaagagaagaagaattactagcagaagaagatgct 701
Db      302  ACTGATTAACCCACTAATCTGATTCATGAGAAAGATCTCAACCGTTAACTGCTGTCT 361
Qy      702  accaatcttagaagaagttgctaaagtaagtaacccctcctctatctatgctgaagct 761
Db      362  GAAACCATTTGAGATTAGCTCTTAAGAGACAAAGCCCTCTCTTGAATTTGGCAAGATGT 421
Qy      762  agaagtggaagcacttgcaacacttgtagtcaataagctcgttggaagcactcgaagttg 821
Db      422  GGACATTAAGACGACCTGCCACCTTATTTGAACAAGCTCCGCTCGAATCAAGGTTTG 481
Qy      822  agcgttaaaagctcctcgtgttctgtagaagcgttaagctatgctggaagatgctat 881
Db      482  TGCATCAAAAGCCAGCGCTTGTGTAAGAAAGAAAGCTTTTGGCAAGATCTTGCCTAT 541
Qy      882  ccttaactggaagaagaacatatattgaagaatgtagtaaaagcttgaagaatgtaagct 941
Db      542  TTTAACTGGAGCCCAAGTATTAACACAGACGCTTGACATGAAATTTTGGAGTT 601
Qy      942  gttcttcttagaagaacagctaaacgttgtagtatttgacaagaagaataactactactg 1001
Db      602  CGAGATGCTGGGAGACATGTAAGAGGCGCACTCTCTTAAGATGATGATCTGATCTTGA 661
Qy      1002  tgggtctggaagaacagaagaatattgaagctcgaggttaacaactctcgccacaatg 1061
Db      662  TGGTGTGTGTGAAGAACTTCATAGAGAGAGGATGAGATGATGATGATGATGATGATGAT 721
Qy      1062  agaaacaagctcagatagatcgtgaaagaactcgaagaagcttgcgaagaactgttg 1121
Db      722  ACAGACACATCTGACTATGACAGGAGAGAGTTCAGAGAAAGCTAGCTAAGCTTTCAGG 781
Qy      1122  tggagt 1127
Db      782  GGTCGT 787

RESULT 10
Bg321293

```



```

Db      25  CTGGCCAGAGATGAGCATGAATCTATCTGATGAAACCAACGACTGAAGACTTGAG 84
Oy      565  gtaglgaagaagaaglttagaccgtggtactctctccataactttagaactaact 624
Db      85  GTTGTGAGGGGTATGAACCTTGATAGGGGATACATCTCTCCCTATTTTCTTAAATATGCC 144
Oy      625  gagaaatagtcttgtaacttgataacctatatactctcttgtaagaaagaatgact 684
Db      145  AAGGACACAAATATGTGAATTTGAGAACCCAGTTATTTCTTATTCAGACAAATAATTTCT 204
Oy      685  agcatgaagaacatgatacctacaaactcttagaacaagtctgaagaacggtccactct 744
Db      205  AGCCTCAATTCGCTGATGATCTTTCTTGAGATGACACTTGAGACAAACCTGCTCTG 264
Y       745  attcttgtagaagacgtagaaggtgaagacacttgtaacacacttgtagaactaagctcgt 804
Db      265  ATTGTTCGGGAGATGTGAGAAAGTGAACCCCTTGCAACACTTGTCTTAATTAATACCTTGA 324
Oy      805  ggaacaccccaagctgtagccgtgaagaagctctgtgttgtagaagcgtgaagaactag 864
Db      325  GCAGGCTTGAGAGATTTGTGGCATATAAAGCCCTGGCTTGGGAGAACGAGAGAGCAAGC 384
Oy      865  ctgaagaataltgatactactcttagagagaagaacaataattgaagatcgtgtataag 924
Db      385  TTGCAAGATCTCCGACACTAATCTGTGGCCAGCTTGAGTGAAGACCTGGGAATGAG 444
Oy      925  ctgaagaatgtaagctgtctctctttagaagaagctaaagctgtagtattgtagaacaag 984
Db      445  TTGGAGAGAGTGACTCTGACATGCTTGTGTAGGCGCAAGCCAGATTAATCTTCGAAAGAT 504
Oy      985  aatactactatcgtttagatggtcgtgaagaacacgaagaatataagctgaagtaagaac 1044
Db      505  GATACCATTAATCTTGATGAGAGCTGGCATTAAGCAAGCTATAGCAAGAAAGTGTAGAG 564
Oy      1045  atcgtgcacaatltgaagaacaagctcagatatagtatcgtgaagaacacctgaagaact 1104
Db      565  ATTGTGAGTATTAATAATCACTCAACATCGATTTATGACAAGAAAGGACAGCAAGCT 624
Oy      1105  ctgcaaaacttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1163
Db      625  CTTGCAAGCTTGTGTGGGGGTGTGTGACATCTTAAGATGTGTGTGTGTGTGTGTGTGTGA 684
Oy      1164  aatgaagaagaagaagaagctgtgtgaagaagctgaagaagaagaagaagaagaagaaga 1223
Db      685  AGTTGGGAGAGAGAAAGATAGAGTACTGATGCTTTGAATGCGACCAAGCGCGAGTTGA 744
Oy      1224  agaagatactgtccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1256
Db      745  GGAGGCTATGTGGTCCANCAGAGAGTGTCCTTT 777

```

RESULT 12

AL564550 899 bp mRNA EST 16-FEB-2001

LOCUS AL564550 LTI.NFL001.NBC4 Homo sapiens cDNA clone CS0DM007Y107 5

DEFINITION prime, mRNA sequence.

ACCESSION AL564550

VERSION AL564550.1 GI:12915068

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 899)

AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.

FEATURES

Location/Qualifiers

```

source
1. 899
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DM007Y107"
/clone_1bp="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com"
BASE COUNT 278 a 176 c 215 g 227 t 3 others
ORIGIN
Query Match 16.7%; Score 275.8; DB 106; Length 899;
Best Local Similarity 59.3%; Pred. No. 1.5e-59;
Matches 464; Conservative 3; Mismatches 316; Indels 0; Caps 0;
Oy      6  ttctaaagaatctcttcttgatgtaagccgtgtaaaacttccagaggtgtagataa 65
Db      116  TGCCAAAGATGTAAATTTGTGGAGAGTCCCGACCTTAATGCTTCAAGGTGAGACCT 175
Oy      66  acttgcaaatgctgttaagaatacaacttgaaacaaagccgtgaatgctgtattgtaaa 125
Db      176  TTTAGCCGATCTGTGGCGCTTACATGTGGGCAAGGAAGAACATGATATTGAGCA 235
Oy      126  gttcttggtcccaagttatcaaaagatggtgtatctgttgcaagaagaatgaaact 185
Db      236  GAGTTGGGAAGTCCCAAGTATACAAAGATGCTGTGCTGTGCAAGTCAATTGACTT 295
Oy      186  tgaagataagcttgtaaaatctggtgcctcaaatgtgttaagaagtagtcccaaaactag 245
Db      296  AAAAGATTAATTAACAGATCATTTGAGAGTAACTTTTCAAGATGTGGCAATTAACAA 355
Oy      246  cgatactgctgtgtgtaagtaacaaacagcaagctcttgcaagaactatattatcgtga 305
Db      356  TGAAGAACCTGGGAGATGACACTACCTACTTACTGTGACCTCTATGTATGCCAAGGA 415
Oy      306  aggtgtaaactgttagcagcgtgtcgttaattctatggccattaaagcttgagatataa 365
Db      416  AGGCTTGAGAGATTTAGCAAGAGTGTATATCACTGGAATCAGAGAGGTGTATGTT 475
Oy      366  agctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 425
Db      476  AGCTTTGATGCTGTATATTTGCTGAACCTTAATAAGAGAGTAACTGTGACCAACCCCTGA 535
Oy      426  agaaatagctcaaglttgaaacacttctgcaaacctctgatacaacaataagtaatacat 485
Db      536  ACAAATTCGACAGTGTGCTACGATTTCTGCAACGAGCAAGAAATATGGCAATATCAT 595
Oy      486  agctgaagctatgctgtaagttgaagaagaggtgtgtatcaagttgaagaagctaaag 545
Db      596  CTCTGATGCAATRAAAAAGTTGAGARAAGGGTGTATCACTACAGTAAGAGATGGAANAAC 655
Oy      546  tctgaaactaatatagatgtgtgtgaagaaglttgacggtgtgtgtgtgtgtgtgtgtgt 605
Db      656  ACTGAATGATGAATTAAGAAATTTAGAGACGATGAAGTGTGATGCGAGCTATATTCTCC 715
Oy      606  atacttgtaactaatcctgagaataatgtgtgtgaactctgataacacctataactcttg 665
Db      716  ATACTTATTAATTAATCAAAAGGTGCAAGATGTGAATCCAGAGTCCCTATGTTCTGT 775
Oy      666  taatgagaaagaattactagcaatgaaagaacatgtaacaaacttgaagaaggtgtctaa 725
Db      776  GAGTGAAAGAAATTTTGTATGATCCAGTCACTTGTACTGCTTGAATATGCCAATGC 835

```

Oy 726 agtaaacccctccactatattatctgtaagacgaagtaagtgtaagcactgcaacact 785
 Db 836 TCACCGTAGCCTTGGTCATATCGCGAAGATGTTGATGAGAGAGCTTAAGTACT 895
 Oy 786 tgt 788
 Db 896 CGT 898

RESULT 13

AL531904

LOCUS AL531904 902 bp mRNA EST 13-FEB-2001
 DEFINITION AL531904 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DM003YJ06 5
 ACCESSION AL531904
 VERSION AL531904.1 GI:12795397
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 902)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 unpublished (2001)
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

COMMENT

Location/Qualifiers
 1..902
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DM003YJ06"
 /clone_1ib="LTI_NFL001_NBC4"
 /sex="male"
 /tissue_type="neuroblastoma cells"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

FEATURES

source

BASE COUNT 278 a 179 c 219 g 225 t 1 others
 ORIGIN

Query Match 16.3%; Score 269; DB 106; Length 902;
 Best Local Similarity 59.5%; Pred. No. 8.1e-58;
 Matches 455; Conservative 0; Mismatches 310; Indels 0; Gaps 0;

Oy 6 tcttaagaacatcccttttgatgctaaagccgtgtaaaacttccgaggtgtaataa 65
 Db 130 TCCCAAGATGTAAATTTGGTGACAGATGCCAGCCTTAATGCTTCAAGGTGTAGACCT 189
 Oy 66 acttgcaaatgctgttaagtaacacttgacctaagaagcgtaatgctgtatgtaaa 125
 Db 190 TTAGCGATGCTGTGGCGGTACATAGGCGCAAGAGACAGTATTTAGCA 249
 Oy 126 gctcttggttcccgatgtaacaaagaatggtatctgttgcacaaagaattgaact 185
 Db 250 GAGTTGGGAAGTCCCAAGATCAACAAAGATGTTGACTGTGCCAAAGTCAATTGACTT 309
 Oy 186 tgaagaatggttgaagaataatggaatggttaagtaagtaagcctcccaaatg 245
 Db 310 AAAAGTAATATACAAAACATTTGAGACTTAAGTCTTCAAGATGTTGCCAATTAACACAA 369

Oy 246 cgaatattgctgtgtaagtaacatacagcaacagtccttgacaaactatattcgtga 305
 Db 370 TGAAGAAAGTGGAGTGGCACTTACCTGCTACTGCTGACCTCTATAGCCAAAGA 429
 Oy 306 aggtgtaaaacttgtaagcagctgtgtaacccatgagcccttaaaagcttgacataaga 365
 Db 430 AGCCTTGAGAGATTTAGCAAGAGTCTTAATCCAGTGAATTCAGAGAGGTGTGATGTT 489
 Oy 366 agctgtgtgtgtgttactaaagaacgaacatlaacaaagcctactcgtgacaaa 425
 Db 490 AGCTGTGATGCTGTATTTGCTGAACCTTAATAAACAGACTTAACCTTGACACCCCTGA 549
 Oy 426 agaaatgcccaagttggaacatttctgcaaacctcgatacaacaatagtaatacat 485
 Db 550 AGAATTTGACAGGTTCTCTGATTTCTGCAAAAGAGCAAAAGAAATTTGGCAATATCAT 609
 Oy 486 agctgaagcctatgctaaagttggaagaagagtgatatacagttgaagaagctaaag 545
 Db 610 CTCGATGCAATGAAAAAGTTGGACAAAGAGGCTCATCATCACTAAAGATGGAAGAAC 669
 Oy 546 tcttgaaactacatlaagatgctgtaaggaagatggaagctgacccctctc 605
 Db 670 ACTGAATGATGAATTAAGAAATTTATGAAAGCATGAAGTTGATGAGAGCTATTTCTCC 729
 Oy 606 atactgtgtacacatccctgaagaatggtgtgtgtaactgtgaacccctataccttg 665
 Db 730 ATACTTATTAATACATCAAAAGCTCAGAAATGTGAATTCAGAGATGCTTATGTTCTGTT 789
 Oy 666 taatgaaaaaagattactgcatgaaagacagtcacacacacttgtaagaagtgctaa 725
 Db 790 GACTGAAAGAAATTTCTGATCCAGTCCATTTGACCTGCTCTTGAATTTGCCAATGC 849
 Oy 726 agtaaacccctccactatattatctgtaagacgaagtaagtgta 770
 Db 850 TCACCGTAGCCTTGGTCATATCGCTGAGATGTTGATGAGCA 894

RESULT 14

AL555787

LOCUS

AL555787

DEFINITION

AL555787 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DM002YF16 5

ACCESSION

AL555787

VERSION

AL555787.1 GI:12897849

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 902)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

unpublished (2001)

CONTACT: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1..902

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0DM002YF16"

/clone_1ib="LTI_NFL006_PL2"

/tissue_type="Placenta"

/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL :

BASE COUNT 277 a 173 c 218 g 230 t 4 others
 ORIGIN

Query Match 16.3% Score 268.2; DB 106; Length 902;
 Best Local Similarity 58.5%; Pred. No. 1.3e-57;
 Matches 459; Conservative 4; Mismatches 322; Indels 0; Gaps 0;

OY 6 tcttaagaatcccttttgatgctaaagccggtgaagaacttccagggtgataaa 65
 117 tcccaaaagatgtaaaatttggtgcagatgcccagccttaagctttaaagtgtaacct 176
 OY 66 acttgaatgctgcttaagaacttgaccctaaagccgtatgcttataaa 125
 177 tttaaccgcatgctgctgcccgttacatggcccaaaagacagacgtattttttagca 236
 OY 126 gtccttggcttcccgatctatcaaaagatggtatcgttgcataaaagaaatgaaact 185
 237 gatttggaagatcccaaaagtaaaagatggtgactgtgcaaaagtcataatgactt 296
 OY 186 tgaagataagcttgaagaataagggcctcaaatggttaagaagatgctcccaaaactg 245
 297 aaaaagtaaatcnaagacattggaacttaactgttcaagatggtgcaaatcaca 356
 OY 246 cgaatgctgctgctgaagacacacagacagcctcctgacacagctatcctgta 305
 357 tgaagaagctggcgatggcacttaccctgctgactgctgacgctcttaccacagca 416
 OY 306 aagtgtaaaacttgaagcagctgctgctgctgctgctgctgctgctgctgctgct 365
 417 agccttgcagaaagattagcaaaagcttaccctgctgctgctgctgctgctgctgct 476
 OY 366 agctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 425
 477 agctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 536
 OY 426 aagaatagctgaagctggaacacatcttgcacacacacacacacacacacacacac 485
 537 agaaattgcac 556
 OY 486 agctgaagctgaagctgaagctgaagctgaagctgaagctgaagctgaagctgaag 545
 597 ccttgatgcaaaagaaagctgaagctgaagctgaagctgaagctgaagctgaagctga 656
 OY 546 tcttgaactacatgatactgctgctgaagaaagaaagctgacccgtgctgctgctgct 605
 657 actgaatgatactgatactgatactgatactgatactgatactgatactgatactg 716
 OY 606 atacttgaactac 665
 717 atacttgaactac 776
 OY 666 taatgaaagaaagattactgaagaaagaaagaaagaaagaaagaaagaaagaaagaa 725
 777 gactgaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaa 836
 OY 726 aagaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaa 785
 837 tcacccgtaagccttggctgatactgctgatactgctgatactgctgatactgctg 896
 OY 786 tgaag 790
 897 tcttg 901

RESULT 15
 LOCUS D46006 716 bp mRNA EST 09-MAR-2000
 DEFINITION R1CS10372A Rice green shoot Oryza sativa cDNA, mRNA sequence.
 ACCESSION D46006 GI:7212768
 VERSION D46006.2
 KEYWORDS EST.

SOURCE
 ORGANISM
 Oryza sativa.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 716)
 Sasaki, T., Miyao, A. and Yamamoto, K.
 Rice cDNA from callus 1995
 JOURNAL
 Unpublished (1995)
 On Mar 9, 1995 this sequence version replaced gi:699715.
 CONTACT: Takuji Sasaki
 National Institute of Agrobiological Resources
 Rice Genome Research Program
 2-1-2 Kanondai, Tsukuba
 Ibaraki,
 Japan 305
 Tel: 0298-38-7441
 Fax: 0298-38-7468
 Email: tsasaki@abr.affrc.go.jp
 PROJECT "RGP"
 Sequence updated (01-Mar-2000).
 Location/Qualifiers

FEATURES
 Source
 1..716
 /organism="Oryza sativa"
 /strain="Nipponbare"
 /db_xref="taxon:4530"
 /clone.lib="Rice green shoot"
 /note="Green shoot (8 days old)"
 Sequence updated (01-Mar-2000).

BASE COUNT 222 a 120 c 198 g 175 t 1 others
 ORIGIN

Query Match 16.2% Score 267.6; DB 156; Length 716;
 Best Local Similarity 61.1%; Pred. No. 1.7e-57;
 Matches 432; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

OY 541 aagagcttgaactacatagatgctgctgaagaaagaaagctgacgtgctacac 600
 4 agaaattgcagaaatattctgctgctgctgctgctgctgctgctgctgctgctgctgct 63
 OY 601 tctcaacttgaactac 660
 64 tcccgctattttgtaactgac 123
 OY 661 ctttgaagaaagaaagattactagcaagaaagaaagaaagaaagaaagaaagaaagaa 720
 124 ctttggctgac 183
 OY 721 gctaagtaaacgctcactccttattatgctgaagaaagaaagaaagaaagaaagaa 780
 184 ataaagctgcaataccac 243
 OY 781 acactgtagcaataagctcctgtagacacacacacacacacacacacacacacacacac 840
 244 actcttggctgcaataagctcctgtagacacacacacacacacacacacacacacacac 303
 OY 841 ttggtgaagcgcgttaagaaagctgctgaagaaagctgctgctgctgctgctgctgct 900
 304 ttggtgtagcgcac 960
 OY 901 atattgaagaaagctgctgaagaaagctgctgaagaaagctgctgctgctgctgctgct 960
 364 atcagagatgaactgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 423
 OY 961 aaacgtgtagtattgaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaa 1020
 424 gcaaaagctgctttaaacaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaa 483
 OY 1021 gatattaaagctgagattaaacaagaaagaaagaaagaaagaaagaaagaaagaaagaa 1080
 484 gaagctgactgaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaa 543
 OY 1081 gatcgtgaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaa 1140

Db 544 GACAAGCAAAAGCTCATGACAGGATTGCAAGCTTCTGCTGTGTCAGTCATTTCAG 603
QY 1141 gtcggagctgctactgaactgaatgaagaagaagaatcgtgtaagaagaatgctcta 1200
Db 604 GTGGAGACACAAACAGAGACCGAACTAAAGAGAGAGAGTTGAGAGTTGAGATCCCTA 663
QY 1201 aatgcacaagaagctgctggtgaagaagatgctcctggtggtg 1247
Db 664 AATGCTACCAAGCTGCTGTGAGAGAGTATTGTTGTTGTTGAGAG 710

Search completed: August 1, 2001, 15:57:12
Job time: 1744 sec